ANALYSIS OF GEOMETRIC SHAPES WITH VARIFOLD REPRESENTATION

by
Hsi-Wei Hsieh

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
June, 2021

Abstract

This thesis is concerned with the theory and applications of varifolds to the representation, approximation, and diffeomorphic registration of shapes. Originating from geometric measure theory, the theory of varifolds provides a convenient way to represent geometric shapes like curves, surfaces or, submanifolds both in continuous and discrete settings. Previous works in shape analysis have made use of this representation as a surrogate to design numerically tractable fidelity terms for curve and surface registration problems. So far, these approaches have primarily focused on processing submanifold data and were not designed to handle more general structures. The varifold representation however provides a very flexible framework that is not restricted to submanifolds but its generality has not yet been exploited to its full extent in shape analysis. In this work, we take a step in this direction by considering deformations acting on general varifolds, and propose a mathematical model for diffeomorphic registration of varifolds under a natural group action which we formulate as an optimal control problem. This new framework allows us to tackle diffeomorphic registration problems for a much wider class of geometric objects and lead to a more versatile algorithmic pipeline.

Varifold matching frameworks heavily rely on the kernel metrics defined on the varifolds spaces. However, the properties of this type of metrics and their relationships with the classical metrics/topologies on measure spaces have not been investigated thoroughly yet. In this work, we study in detail the construction of kernel metrics on the space of varifold and the resulting topological properties of those metrics. Based on

these results, we address the problem of optimal finite approximations (quantization) for kernel metrics, propose a projection-based approach for varifold representation, and show a Γ-convergence property for the discrete registration functionals.

In the last part of this thesis, we tackle the imbalanced shape matching problems, namely the situation in which the source and target shapes involve considerable variations of mass or density which cannot be entirely described by diffeomorphic transformations. We extend our varifold matching model by augmenting the diffeomorphic component with a global or local density changes. Based on the optimality conditions provided by the Pontryagin maximum principle, we derive a shooting algorithm to numerically estimate solutions and illustrate the practical interest of this model for several types of geometric data such as fiber bundles with inconsistent fiber densities or partially observed and incomplete surfaces.

Thesis Readers

Dr. Nicolas Charon (Primary Advisor)
Assistant Professor
Department of Applied Mathematics and Statistics
Johns Hopkins University

Dr. Laurent Younes
Professor
Department of Applied Mathematics and Statistics
Johns Hopkins University

Dedicated to my parents for their unconditional love and support.

Acknowledgements

First of all, I would like to express my gratitude and appreciation to my advisor, Prof. Nicolas Charon, for introducing me to the field of shape analysis and for his guidance in the past six years. He is always very patient, supportive, and generous in sharing his knowledge and research experience. His passion and dedication for applied math research deeply inspired me, and I feel very lucky to have him as my advisor during my Ph.D. study.

Secondly, I would like to thank Prof. Laurent Younes for serving as my thesis reader and committee member. I learn the fundamental knowledge of shape analysis from his inspiring course mathematical foundations of computational anatomy and his book shape and diffeomorphisms. I also want to thank Prof. Chikako Mese for serving as my committee member and introducing me to many interesting topics in geometric analysis. Her topic course about the geometric variational problems and harmonic maps between manifolds is my favorite course at Hopkins. Moreover, I would like to thank Prof. Carey Priebe, Prof. Mario Micheli, for serving in my committee.

I would like to thank my co-authors and collaborators, including Prof. Martin Bauer, Prof. Philipp Harms, Prof. Archana Venkataraman, Ravi Shankar, Dr. James Lu, and Dr. Kai Liu, for their patient discussion and support. These collaborations broadened my vision of scientific research. Special thanks to Dr. Wen-Kai You and Dr. You-Ping Yang for helping me improving my presentation skills.

I also want to thank my dearest friends in CIS and department of Applied Mathe-

matics and Statistics: Gongkai Li, Wei-Chun Hung, Chu-Chi Lee, Ting Chao, Yashil Sukurdeep, Ke Chen, Kamel Lahouel, Sue Kulason, Chin-Fu Liu, and Michele Bierbaum Lohr for the delightful days at Hopkins.

Finally, I would like to thank my parents and my wife for their support and encouragement.

Contents

List of Figures

Chapter 1

Introduction

The work presented in this thesis belongs to the field of *Computational Anatomy*, as introduced in the seminal work of Grenander and Miller [1]. The main objective of computational anatomy is to design mathematical models and numerical algorithms to provide theoretical foundations for statistical analysis of anatomical shape data sets. Typical statistical problems of interest include: (1) *analysis of populations*: estimate mean shape and morphological variability [2–5]; (2) *classification/discrimination*: classify pathologies from morphological deviations [6–9] (3) *learn the temporal evolution*: descriptive or predictive model of organ development across time [10–13].

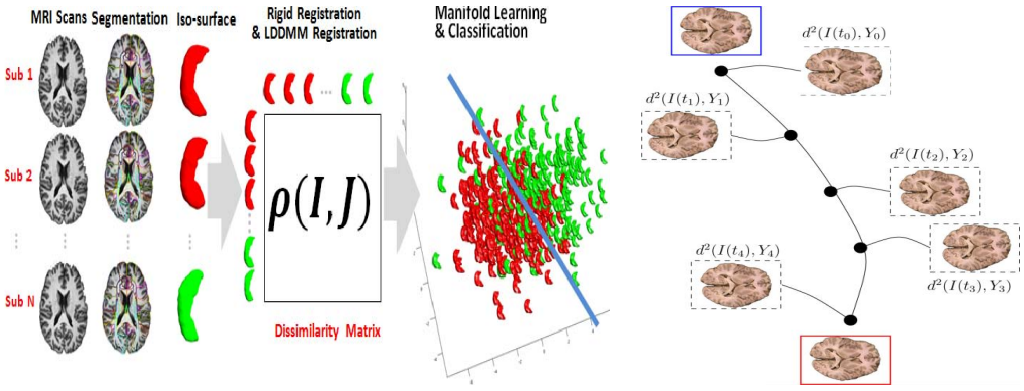
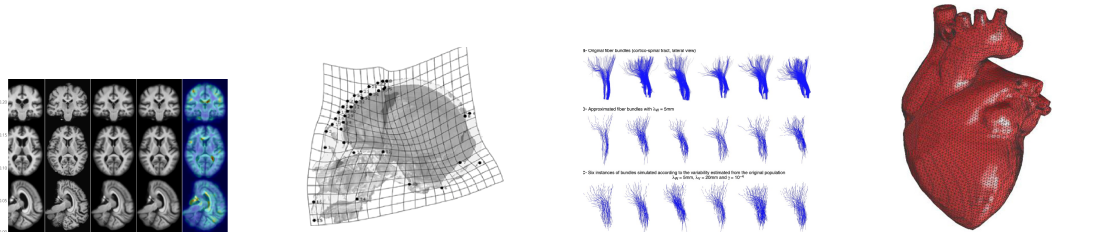


Figure 1-1. Shape classification [7] and geodesic regression [11]

The development of such statistical shape analysis benefits from the advances of non-invasive acquisition technologies (e.g. magnetic resonance imaging (MRI),

diffusion tensor imaging (DTI)) which can generate a large amount of anatomical data. Nevertheless, the development of these technologies also came with challenges. Firstly, new technologies result in numerous modalities of shape data types to be analyzed, going from landmark points sets, curves, surfaces, brain fiber sets, biomedical images, etc (See figure 1-2). Secondly, the fundamental nonlinearity that exists among almost all shape data types encountered in applications. Such a lack of linearity makes even the most basic statistics like means, median, or covariance challenging to define and compute. In his address to the ICM in 2002 [14], David Mumford even described shapes as the "ultimate nonlinear sort of thing". In computational anatomy, the central idea to treat these difficulties is to introduce the notion of *shape space* from differential geometry. Under such notion, one can define distances between shapes, and from which means and shape variability can be estimated from population shape data. Moreover, if a shape space can be endowed with a Riemannian structure, one can further introduce the concept of energy minimizing paths on the shape space, which is the foundation for regression models and longitudinal studies mentioned in the previous paragraphs. We shall briefly present the principle of building shape spaces in the following section.



Images [15] Landmark points [16] Brain fibers [4] Meshed surface [17]

Figure 1-2. Shape Data

1.1 General principle of shape analysis

The underlying principle of shape space is to compare variability between two shapes through quantifying the deformation needed to warp one shape to the other. Such an idea was first pioneered by D'Arcy Thompson [18] and was formulated mathematically by Grenander in his theory of deformable templates [19, 20]. To be rigorous, in this point of view, a shape space is a homogeneous G-space M, which is defined as a set endowed with a transitive left action by a group G of transformations in $\mathbb{R}^n$. The transitivity of the action means that any pair of shapes m_0 and m_1 in M can be joined by a transformation, i.e. $\exists g \in G$ such that $g \cdot m_0 = m_1$. Or equivalently, any shape in the shape space can be reached by transformations of a fixed *template* shape m_0, that is, the shape space is the orbit of m_0, $G \cdot m_0 = M$. As an simple example, we can take $M = G = \mathbb{R}^n$ and the action defined as the translations: $g \cdot m \doteq g + m,\ \forall g, m \in \mathbb{R}^n$. However, finite dimensional groups such as translations or rotations are not rich enough since one usually needs the deformation to perform complex and localized changes in shapes. Therefore, we usually choose the deformation group G to be a subgroup of $\mathrm{Diff}(\mathbb{R}^n)$, the group of C^1-*diffeomorphisms* of $\mathbb{R}^n$. A *diffeomorphism* $\varphi : \mathbb{R}^n \mapsto \mathbb{R}^n$ is an invertible smooth map with smooth inverse. The advantages of choosing this type of maps are that diffeomorphisms prevent folding and preserve the topological and smooth structure of geometric objects. The followings are three examples of actions of deformation groups on shape spaces.

Example 1 (Labelled points.)**.** *We consider the the space of all distinct d landmark points in $\mathbb{R}^n$,*

$$M = \{m = (x_1, \cdots, x_d) \in \mathbb{R}^{nd} \mid x_i \neq x_j,\ \forall\ i \neq j\}.$$

The left action of the deformation group $G = Diff(\mathbb{R}^n)$ is defined as

$$g \cdot m \doteq (g(x_1), \cdots, g(x_d)), \forall g \in G,\ \forall m \in M.$$

Example 2 (Images.)**.** *We consider images to be square integrable functions defined on a fixed domain $\Omega \subset \mathbb{R}^n$, i.e. $L^2(\Omega, \mathbb{R})$. One can define a left action of $G = Diff(\mathbb{R}^n)$ on $L^2(\Omega, \mathbb{R})$ by*

$$g \cdot I \doteq I \circ g^{-1}, \ \forall g \in G, \ \forall I \in L^2(\Omega, \mathbb{R}).$$

For any fixed image $I_0 \in L^2(\Omega, \mathbb{R})$, $M \doteq G \cdot I_0$ is a homogeneous space with the action defined above.

Example 3 (Measures.)**.** *Let μ_0 be a Borel measure in $\mathbb{R}^n$ and consider the pushforward of measures defined by*

$$(g \cdot \mu_0)(A) \doteq \mu(g^{-1}(A)), \forall \ g \in G = \mathit{Diff}(\mathbb{R}^n),$$

*and for all Borel set $A \subset \mathbb{R}^n$. Then $M \doteq G \cdot \mu_0$ is a homogeneous space with the transitive action defined by the pushforward of measures. One important case is that an unlabelled set of points $X = \{x_1, \cdots, x_N\}$ in $\mathbb{R}^n$ can be represented as a discrete measure by its **empirical measure**: $\mu_X = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$. In this case, the measure pushforward commutes with the action on points defined in example 1, namely:*

$$\varphi \cdot \mu_X = \frac{1}{N} \sum_{i=1}^{N} \delta_{\varphi(x_i)} = \mu_{\varphi(X)}.$$

Now assuming that G is equipped with a right-equivariant distance d_G, i.e. $d_G(g_0 g, g_1 g) = d_G(g_0, g_1)$, for all $g, g_0, g_1 \in G$. Given two shapes m_0 and m_1 in a shape space M and a deformation $g \in G$ such that $g \cdot m_0 = m_1$, the quantity $d_G(Id, g)$ can be interpreted as the cost of deforming the shape m_0 to m_1 by the deformation g. One can then compare m_0 and m_1 by finding the minimal cost and the optimal deformation to transport m_0 to m_1, which is called *shape registration*. Moreover, the minimal cost obtained from the shape registration process defines a pseudo-distance on the shape space M, which is stated as the following Theorem from Chapter 11 in [21]:

Theorem 1. *Assume that G acts transitively on M and d_G is a right-equivariant distance on G, then*

$$d_M(m_0, m_1) \doteq \inf_{g \in G} \{d_G(id, g) \mid g \cdot m_0 = m_1\}$$

defines a pseudo-distance on M. Moreover, if the isotropy group $G_{m_0} \doteq \{g \in G | g \cdot m_0 = m_0\}$ is closed, then d_M is a distance.

This approach has the advantage of shifting the problem of metric construction from the many different cases of shape spaces to the single setting of deformation groups. As stated in Theorem 1, the fundamental requirement is the right-invariance of the metrics on the deformation groups. To build the right invariance metrics on subgroups of $\mathrm{Diff}(\mathbb{R}^n)$, we first assume in addition that G is a finite dimensional Lie group (e.g. groups of affine maps or isometries), then from a classical result in Riemannian Geometry, one can construct right invariant Riemannian distances on G (cf. Chapter 3 in [22] or Chapter 1 in [23]). Indeed, one can define an inner product $\langle \cdot, \cdot, \rangle_{id}$ on the Lie algebra of G, $T_{id}G$, then extend this inner product to a Riemannian metric on G by pullback of $\langle \cdot, \cdot, \rangle_{id}$ to each point $g \in G$ through the inverse of right translations:

$$\langle u, v \rangle_g \doteq \langle d_g R_{g^{-1}}(u), d_g R_{g^{-1}}(v) \rangle_{id}, \tag{1.1}$$

where $R_g(g') \doteq g'g$ is the right translation and $d_{g'}R_g$ denote the differential of R_g at $g' \in G$. The Riemannian distance defined as

$$d_G(g, g') \doteq \inf_{g_0 = g, g_1 = g'} \int_0^1 |\dot{g}_t|^2_{g_t} dt$$

is a right invariant distance on G. However, finite dimensional Lie groups are not rich enough to describe morphing processes of shapes, thus infinite dimensional subgroups of $\mathrm{Diff}(\mathbb{R}^n)$ are often considered. The Riemannian geometry of infinite dimensional groups of diffeomorphisms has been extensively studied in the setting of nonlinear global analysis [24, 25] and geometric mechanics [26].

The *Large Deformation Diffeomorphic Metric Mapping (LDDMM)* model pioneered in [21, 27] is one such framework that defines Riemannian metrics for diffeomorphic mappings obtained as flows of time-dependent vector fields. Since we'll rely on this framework throughout this thesis, we shall briefly present the LDDMM model in the next section.

1.2 Large deformation diffeomorphic metric mapping model

The purpose of this section is to briefly recap the LDDMM framework for building groups of diffeomorphisms, their right invariant Riemannian metric, and algorithms for shape registrations. In this setting, the groups of diffeomorphisms are constructed by integrating time dependent vector fields [28, 29] and registering two shapes can be generically formulated as an *optimal control* problem [30], the functionals to optimize being typically a combination of a deformation regularization term given by the LDDMM metric on the group and a fidelity term that enforces matching between the two shape objects. Applications of this model have been widespread in particular within the field of computational anatomy, due to the ability to adapt it to various data structures. See e.g. [31, 32] for recent reviews.

1.2.1 Construction of diffeomorphism from flows of time dependent vector fields

The principle of LDDMM is to consider diffeomorphisms constructed from flow maps of time dependent vector fields. Let $v = (v_t)_{t\in[0,1]}$ be a time dependent vector field in $\mathbb{R}^n$, then the *flow* map φ_t^v of v is defined as the solution of the following ODE:

$$\begin{cases} \dot{\varphi}_t = v_t \circ \varphi_t \\ \varphi_0 = \mathrm{id} \end{cases}$$

for $t \in [0,1]$. A more explicit way to express the solution of this ODE is as follows:

$$\varphi_t(x) = x + \int_0^t v_s(\varphi_s(x))ds, \ \forall (t,x) \in [0,1] \times \mathbb{R}^n.$$

Some sufficient conditions can be imposed on the vector field v so that the above ODE admits a solution φ_t^v and φ_t^v is a diffeomorphism for each $t \in [0,1]$. Let us start by considering vector fields in space $C_0^1(\mathbb{R}^n, \mathbb{R}^n)$, the space of C^1 vector fields u such that u and du vanish at infinity and equipped with the norm

$$\|u\|_{1,\infty} \doteq \|u\|_\infty + \|du\|_\infty,$$

where $\|u\|_\infty \doteq \sup_{x \in \mathbb{R}^n} |u|$, $\|du\|_\infty \doteq \sup_{x \in \mathbb{R}^n} \|d_x u\|_{op}$ and $\|d_x u\|_{op} \doteq \sup_{|h| \leq 1} \frac{|d_x u(h)|}{|h|}$ denotes the operator norm. Now the fundamental result for generating diffeomorphisms from ODEs is stated as follows:

Theorem 2. *Assuming the time dependent vector field v satisfies:*

- $v_t \in C_0^1(\mathbb{R}^n, \mathbb{R}^n), \ \forall t \in [0,1]$
- $\int_0^1 \|v_t\|_{1,\infty} dt < \infty$

then the flow map φ_t^v of v is well defined and a diffeomorphism for all $t \in [0,1]$.

To build a Riemannian structure on diffeomorphisms generated from flows, we need to define a Hilbert structure in the space of vector fields. For that purpose, we further assume that the vector field v_t belongs to a Hilbert space which is continuously embedded in $C_0^1(\mathbb{R}^n, \mathbb{R}^n)$:

Definition 1. *Let V be a Hilbert space of vector fields in $\mathbb{R}^n$. We call V an admissible space if $V \hookrightarrow C_0^1(\mathbb{R}^n, \mathbb{R}^n)$, i.e., $\exists c > 0$ such that*

$$\|u\|_{1,\infty} \leq c\|u\|_V, \ \forall u \in V.$$

Given such an admissible Hilbert space V, we denote by $L^2([0,1],V)$ the space of time dependent vector field v such that $v_t \in V,\ \forall t \in [0,1]$ and $\int_0^1 \|v_t\|_V^2 dt < \infty$. It is straightforward to see that all $v \in L^2([0,1,V])$ satisfy the conditions of Theorem 2, hence flows obtained from such vector fields are diffeomorphisms. We define

$$G_V \doteq \{\varphi_1^v \mid v \in L^2([0,1],V)\},$$

the set of all diffeomorphisms obtained as flows of vector fields in $L^2([0,1],V)$. It can be shown [21] that G_V is a group. Analogous to the finite dimensional Lie group case discussed in section 1.1, a right equivariant distance can be defined on G_V. Indeed, V can be formally viewed as the Lie algebra $T_{id}G_V$, which is equipped with the inner product $\langle \cdot, \cdot \rangle_V$. Next, we can extend this inner product to each $\varphi \in G_V$ in order to define a right equivariant metric as follows:

$$\langle u, v \rangle_\varphi \doteq \langle u \circ \varphi^{-1}, v \circ \varphi^{-1} \rangle_V.$$

This is exactly (1.1) with the right translation $R_\varphi(\psi) = \psi \circ \varphi$ and its differential $d_\psi R_\varphi(v) = v \circ \varphi$. Let $\{\psi_t\}_{t\in[0,1]}$ be a path joining $\psi, \psi' \in G_V$. Assuming ψ_t' exists for almost every $t \in [0,1]$, then we can define the the energy of this path as $\int_0^1 \|\dot{\psi}_t\|_{\psi_t}^2 dt$, which can be interpreted as the cost needed to move from ψ to ψ'. Due to the right invariance of the metric, it is equivalent to consider the path $\{\varphi_t^v\}_{t\in[0,1]}$ as a flow of v with $\psi' = \varphi_1^v \circ \psi$ and its energy $\mathcal{E}(v) \doteq \int_0^1 \|v_t\|_V^2 dt$ parametrized by the vector field $v \in L^2([0,1],V)$. The Riemannian distance on G_V can be defined as the minimal energy needed to move from one diffeomorphism to the other:

Theorem 3. *Define*

$$d_{G_V}(\psi, \psi') \doteq \inf_{v \in L^2([0,1],V)} \{\sqrt{\mathcal{E}(v)} \mid \psi' = \varphi_1^v \circ \psi\},\ \forall \psi, \psi' \in G_V.$$

Then d_{G_V} is a right equivariant distance on G_V and (G_V, d_{G_V}) is a complete metric space. Moreover, the above infimum can be achieved by a minimizer $v \in L^2([0,1],V)$.

This result is proved by Theorem 7.17 in [21]. Now with this right equivariant metric d_{G_V}, we can compare shapes based on Grenander's framework discussed in section 1.1. Assuming the shape space M is a homogeneous G_V space, then from Theorem 1, d_{G_V} induces a distance on M which can be written as:

$$d_M(m_0, m_1) = \inf_{v \in L^2([0,1],V)} \left\{ \sqrt{\mathcal{E}(v)} \mid \varphi_1^v \cdot m_0 = m_1 \right\}.$$

The above distance involves an optimization problem which searches for optimal deformations parametrized by vector fields in $L^2([0,1], V)$, mapping the shape m_0 onto m_1. This is called the **exact matching problem** between m_0 and m_1.

We have seen that diffeomorphisms can be generated from ODEs. However, to develop computational methods for flows of diffeomorphisms, we need to choose admissible Hilbert spaces so that vector fields can be expressed in more explicit forms. For that purpose, we choose V to be a reproducing kernel Hilbert space (RKHS) of smooth vector fields. The notion of RKHS was originally introduced in functional analysis by Aronszajn [33] and has wide applications in statistics and machine learning [34, 35]. Let us recap some important definitions and properties of RKHS.

Definition 2. *Let $(V, \langle \cdot, \cdot \rangle_V)$ be a Hilbert space of $\mathbb{R}^\ell$ valued functions on a set Ω. We call V a reproducing kernel Hilbert space if the linear functional $\delta_x^\alpha : f \in V \mapsto \alpha \cdot f(x)$ continuous for all $x \in \Omega$ and $\alpha \in \mathbb{R}^\ell$.*

We denote by $\mathbf{K}_V : V^* \to V$ the Riez isometry and its inverse $\mathbf{L}_V = \mathbf{K}_V^{-1} : V \to V^*$. V has the following *reproducing property*:

$$\alpha \cdot f(x) = (\delta_x^\alpha | f) = \langle \mathbf{K}_V \delta_x^\alpha, f \rangle_V.$$

For all $y \in \Omega$, the map $\alpha \mapsto (\mathbf{K}_V \delta_x^\alpha)(y) \in \mathbb{R}^\ell$ is linear in α, hence there exists a matrix valued function $k_V(\cdot, \cdot) : \Omega \times \Omega \mapsto \mathbb{R}^{\ell \times \ell}$ such that $(\mathbf{K}_V \delta_x^\alpha)(\cdot) = k_V(x, \cdot)\alpha \in V,\ \forall x \in \Omega,\ \alpha \in \mathbb{R}^\ell$ and

$$\langle k_V(x, \cdot)\alpha, k_V(y, \cdot)\beta \rangle_V = \alpha \cdot (k_V(x, y)\beta),$$

$\forall x, y \in \omega$ and $\alpha, \beta \in \mathbb{R}^\ell$. It is straightforward to verify that the function $k_V(\cdot, \cdot)$, we call the *reproducing kernel* of V, is positive definite. The definition of a positive definite kernel is stated as follows:

Definition 3. *We say that the function $k : \Omega \times \Omega \to \mathbb{R}^{\ell \times \ell}$ is a positive definite kernel on the set Ω with values in $\mathbb{R}^\ell$ if it satisfies*

- $k(x, y) = k(y, x)^T, \ \forall x, y \in \mathbb{R}^\ell$
- *For $N \in \mathbb{N}$, $\alpha_1, \cdot, \alpha_N \in \mathbb{R}^\ell$ and $x_1, \cdot, x_N \in \Omega$, we have*

$$\sum_{i,j=1}^{N} \alpha_i \cdot (k(x_i, x_j)\alpha_j) \geq 0.$$

Conversely, given any reproducing kernel k, it was proved by Aronszajn [33] that there is a unique RKHS V whose reproducing kernel is k:

Theorem 4. *Given a positive definite kernel $k : \Omega \times \Omega \mapsto \mathbb{R}^{\ell \times \ell}$, there exists a unique RKHS of functions from Ω to $\mathbb{R}^\ell$ that has k as its reproducing kernel.*

Let us now come back to the problem of constructing admissible spaces for generating diffeomorphisms. We assume, from now on, that the kernels are defined on an *open, connected* subset Ω of $\mathbb{R}^m$. With Theorem 4, we can obtain RKHS of vector fields in $\mathbb{R}^n$ by specifying a kernel on $k : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^n$. However, to ensure that the corresponding RKHS are admissible, we need some regularity conditions on their kernels. To state the result, let us first introduce the multiindex notation. A vector of the form $p = (p_1, \cdots, p_r)$, where each component p_i is a nonnegative integer, is called a multiindex of order $|p| = p_1 + \cdots + p_r$. Given a multiindex p, we denote by

$$\partial^p v(x) = \frac{\partial^{|p|}}{\partial_{x_1}^{p_1} \cdots \partial_{x_r}^{p_r}} v(x)$$

the derivative of the vector field $v \in C^{|p|}(\Omega, \mathbb{R}^\ell)$ with respect to the multiindex p. For a kernel function k, we denote by $\partial_i^p k$ the partial derivative of k where ∂^p is taken to

i-th variable of k. Then the classical result is stated as the following Theorem, which can be found in [36]:

Theorem 5. *Let V be a RKHS with kernel $k_V : \Omega \times \Omega \mapsto \mathbb{R}^{\ell\times\ell}$ and $s \geq 0$ be an integer, then the following two statements are equivalent:*

- $V \hookrightarrow C^s(\Omega, \mathbb{R}^\ell)$
- *$\partial_1^p \partial_2^p k_V$ exists for all multi-indices p with $0 \leq |p| \leq s$, is continuous in each of the two variables (separately), and is locally bounded.*

Under above assumptions, the following also holds:

- *For all $(x, \alpha) \in \Omega \times \mathbb{R}^\ell$, and multi-indices p with $0 \leq |p| \leq s$, we have $\partial_1^p k_V(x, \cdot)\alpha \in V$ and*

$$\langle \partial_1^p k_V(x, \cdot)\alpha, u\rangle_V = \alpha \cdot \partial^p u. \tag{1.2}$$

If furthermore, $k_V(x, \cdot)\alpha \in C_0^s(\Omega, \mathbb{R}^\ell)$ $\forall (x, \alpha) \in \Omega \times \mathbb{R}^\ell$, then $V \hookrightarrow C_0^s(\Omega, \mathbb{R}^\ell)$.

If in particular, a kernel $k_V : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^n$ satisfies the condition in the above Theorem with $s = 2$, then the corresponding RKHS V is an admissible space. Equation (1.2) is called the *derivative reproducing property*, which was originally proved in [37].

Now with Theorem 4 and Theorem 5, we can build admissible space from a smooth kernel. With this approach, *radial* kernel is often chosen to ensure rotation and translation invariance. We call a kernel k a radial kernel if it has the following form:

$$k(x, y) = \rho(|x - y|^2)\mathrm{id}_{\mathbb{R}^n},$$

with $\rho : \mathbb{R} \mapsto \mathbb{R}$. In order to obtain admissible spaces, we can choose radial kernels satisfying $\rho \in C_0^s(\mathbb{R}, \mathbb{R})$, for some $s \geq 1$, then it's clear from Theorem 5 that associated

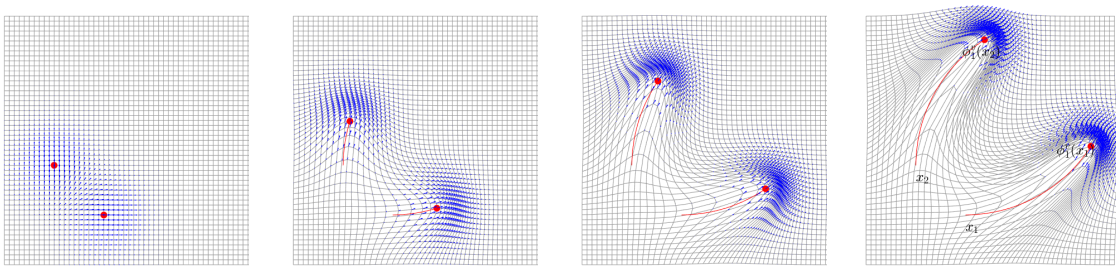
RKHSs are continuously embedded in $C_0^1(\mathbb{R}^n, \mathbb{R}^n)$ as well. The Cauchy kernel given by

$$k(x, y) = \frac{1}{1 + \frac{|x-y|^2}{\sigma^2}} \mathrm{id}$$

is one example of such radial kernel. The Gaussisan kernel

$$k(x, y) = e^{-\frac{|x-y|^2}{\sigma^2}} \mathrm{id}$$

is another such kernel and is widely used in applications. The parameter σ appears in these two kernels controls the scale at which two points interact with each other. In this thesis, we'll mainly rely on the Gaussian kernel to build admissible space of vector fields to generate flows of diffeomorphisms. Figure 1-3 illustrates a path of diffeomorphisms generated by this approach.



$t = 0$ $t = 1/3$ $t = 2/3$ $t = 1.$

Figure 1-3. Flow of diffeomorphisms generated from an ODE with Gaussian kernel.

1.2.2 Inexact matching problem

We have discussed in the previous sections how to build shape spaces in Grenander's framework with groups of diffeomorphisms obtained through flows of vector fields and acting on shape objects. In such a framework, the most important assumption is that the action of the deformation group is transitive on the shapes. However, this assumption is often not reasonable in practical situations. First, the groups of

diffeomorphisms built from admissible spaces are only subsets of $\mathrm{Diff}(\mathbb{R}^n)$, which do not contain all diffeomorphisms in $\mathbb{R}^n$. Second, shapes may not be diffeomorphic to each other as a result of noise, discretization, or reconstruction process. Due to these reasons, the exact registration setting which aims to interpolate between the two shapes by a path of diffeomorphisms may not be feasible and thereby is generally replaced by the following relaxed problem:

$$\inf_{v \in L^2([0,1],V)} E(v) \doteq \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + A(\varphi_1^v \cdot m_0, m_1). \tag{1.3}$$

The second term $A(\cdot,\cdot)$ in the above expression is called the **data attachment term** or the **fidelity term**, which measures the dissimilarity between the deformed shape $\varphi_1^v \cdot m_0$ and the target shape m_1. Unlike the exact matching problem, the minimization problem (1.3) drives the registration by penalizing the dissimilarity between the deformed source and target shapes. The first term $\int_0^1 \|v_t\|_V^2 dt$ plays the role of a regularization term which regularizes the deformation by the energy along the path $t \mapsto \varphi_t^v$.

One fundamental question is the existence of solutions to the minimization problem (1.3). It is shown in [21] Chapter 10 that, with some continuity condition on the data attachment term, there exists a global minimum. However, the uniqueness of the minimizer is not guaranteed in general. The result of existence is precisely the following:

Theorem 6. *If A is bounded from below and the functional $v \mapsto A(\varphi_1^v \cdot m_0, m_1)$ is continuous with respect to the weak topology of $L^2([0,1],V)$ for all $m_0, m_1 \in M$, then there exists a minimizer for the minimization problem (1.3).*

Now with the result of existence, we are ready to move forward to discuss the optimality conditions and develop numerical methods from those conditions. Problem (1.3) can be regarded as an optimal control problem with controls $v \in L^2([0,1],V)$ and the state equation $q_t = \varphi_t^v \cdot m_0,\ t \in [0,1]$. In standard finite-dimensional optimal

control problems, these are provided by the Pontryagin Maximum Principle (PMP) introduced originally in [38]. The approach generalizes, with a certain number of technicalities, to a broad class of infinite-dimensional shape matching problems, as developed in [30, 39–41]. To simplify our presentation, we follow the setting of [30, 32], which assume that the shape space M is an open subset of a Banach space $\mathcal{Q}$ called the *state space*. For more general setting, we refer to the reader [39–41].

In order to introduce the PMP, we need to rewrite the state equation into the form of an ODE. Let us abuse the notation a little and define the mapping $R_q : G_V \mapsto M$ by $\varphi \mapsto \varphi \cdot q$ for any fixed $q \in M$. Its differential is denoted by $\xi_q v \doteq d_{\mathrm{id}} R_q : T_{\mathrm{id}} G_V \mapsto \mathcal{Q}$ and is called the *infinitesimal action* of the vector field v on the state q. Then one can formally verify as a result of chain rule that the evolution of state satisfies:

$$\dot{q}_t = d_{\varphi_t^v} R_q \circ d_{\mathrm{id}} R_{\varphi_t^v}(v_t) = d_{\mathrm{id}} R_{q_t}(v_t) = \xi_{q_t} v_t.$$

Thus problem (1.3) can be rewritten as the following optimal control problem:

$$\begin{aligned} &\dot{q}_t = \xi_{q_t} v_t, \;\; q_0 = m_0 \\ &\min_{v \in L^2([0,1],V)} E(v) = \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + U(q_1), \end{aligned} \tag{1.4}$$

where $U(\cdot) \doteq A(\cdot, m_1)$. We introduce the *Hamiltonian* $H : \mathcal{Q}^* \times M \times V \mapsto \mathbb{R}$ in which $p \in \mathcal{Q}^*$ is the *costate* or the *generalized momentum*:

$$H(p, q, v) \doteq (p | \xi_q v) - \frac{1}{2} \|v\|_V^2. \tag{1.5}$$

We denote by $H^1([0,1], \mathcal{Q}^*)$ the Sobolev space of elements in $L^2([0,1], \mathcal{Q}^*)$ having a weak derivative in $L^2([0,1], \mathcal{Q}^*)$, then the PMP is stated as follows:

Theorem 7. *If v is an optimal control of problem (1.4), there exists a costate p_t in $H^1([0,1], \mathcal{Q}^*)$ such that the following system is satisfied:*

$$\begin{cases} \dot{q}_t = \frac{\partial}{\partial p} H(p_t, q_t, v_t) \\ \dot{p}_t = -\frac{\partial}{\partial q} H(p_t, q_t, v_t) \\ \frac{\partial}{\partial v} H(p_t, q_t, v_t) = 0, \end{cases} \tag{1.6}$$

with the end point condition

$$p_1 + d_{q_1} U = 0.$$

From the *Hamiltonian equation* (1.6), we can now derive further characterizations of the optimal trajectories. First of all, we observe that the Hamiltonian is a constant along such trajectories:

$$\frac{d}{dt} H(p_t, q_t, v_t) = \frac{\partial}{\partial p} H(p_t, q_t, v_t)\dot{p}_t + \frac{\partial}{\partial q} H(p_t, q_t, v_t)\dot{q}_t + \frac{\partial}{\partial v} H(p_t, q_t, v_t)\frac{\partial v_t}{\partial t} = 0.$$

Second, we have $\frac{\partial}{\partial v} H(p, q, v) = \xi_q^* p - \langle \cdot, v \rangle_V$ and the condition $\frac{\partial}{\partial v} H(p_t, q_t, v_t) = 0$ is equivalent to $v_t = \mathbf{K}_V \xi_{q_t}^* p_t$, where $\xi_q^* : \mathcal{Q}^* \mapsto V^*$ is the adjoint operator of ξ_q. Plugging $v = K_V \xi_q^* p$ into the Hamiltonian, we define the *reduced Hamiltonian* $H_r : \mathcal{Q}^* \times M \mapsto \mathbb{R}$ by

$$H_r(p, q) \doteq H(p, q, \mathbf{K}_V \xi_q^* p) = \frac{1}{2}\|v\|_V^2 = \frac{1}{2}\|\mathbf{K}_V \xi_q^* p\|_V^2 = \frac{1}{2}(p|\xi_q \mathbf{K}_V \xi_q^* p).$$

From above and the fact that the Hamiltonian is a constant along the optimal path given by PMP, we also have that $t \mapsto \|v_t\|_V^2$ is a constant. It also straightforward to check from $\frac{\partial H}{\partial v} = 0$ and the chain rule that $\frac{\partial H}{\partial p} = \frac{\partial H_r}{\partial p}$ and $\frac{\partial H}{\partial q} = \frac{\partial H_r}{\partial q}$. We summarize the discussion above with the following result:

Proposition 1. *Let (q_t, v_t) be a solution of problem (1.4). There exists a costate p_t such that $v_t = \mathbf{K}_V \xi_{q_t}^* p_t$, p_t satisfies $p_1 + d_{q_1} U = 0$ and the* ***geodesic equations****:*

$$\begin{cases} \dot{q}_t = \frac{\partial}{\partial p} H_r(p_t, q_t) = \xi_{q_t} \mathbf{K}_V \xi_{q_t}^* p_t \\ \dot{p}_t = -\frac{\partial}{\partial q} H_r(p_t, q_t) = -\frac{\partial}{\partial q} \frac{1}{2}(p_0|\xi_{q_0} \mathbf{K}_V \xi_{q_0}^* p_0). \end{cases} \tag{1.7}$$

Moreover, $t \mapsto \|v_t\|_V^2$ is constant and the cost function can be written as:

$$E(v) = \frac{1}{2}\|v_0\|_V^2 + U(q_1),$$

or equivalently as a function of the initial momentum p_0:

$$E(p_0) = \frac{1}{2}(p_0|\xi_{q_0} \mathbf{K}_V \xi_{q_0}^* p_0) + U(q_1). \tag{1.8}$$

Now we're ready to discuss the numerical methods for problem (1.4) based on Proposition 1. To simplify our discussion, we assume that $\mathcal{Q}$ is finite dimensional. Proposition 1 suggests that, to solve problem (1.4), we can only search for the solution among *geodesics* given by (1.7). Since the dynamics of the state and the optimal control now are determined entirely by the initial costate p_0, the cost functional E can be written as a functional of p_0. We can then solve problem (1.4) by minimize the cost E directly over the initial momentum p_0. This is known as the *geodesic shooting* scheme [30, 42–45] which has been studied extensively in the literature. To minimize E over p_0, we can apply gradient based optimization algorithms such as gradient descent or BFGS. The gradient of $\frac{1}{2}(p_0|\xi_{q_0}\mathbf{K}_V\xi_{q_0}^* p_0)$ with respect to p_0 is simply $\xi_{q_0}\mathbf{K}_V\xi_{q_0}^* p_0$. The gradient of the fidelity term is obtained by first computing the gradient of the fidelity term with respect to q_1, $\nabla_{q_1} U$ then propagate it back to time 0 by solving the linearized adjoint Hamiltonian system backward in time. To be concrete, let's denote $F(q,p) \doteq (\partial_p H_r(p,q), -\partial_q H_r(p,q))$ and introduce the adjoint variables $Z = (\tilde{q}, \tilde{p}) \in \mathcal{Q}^* \times \mathcal{Q}$. Solving the following linearized adjoint Hamiltonian system

$$\begin{cases} \dot{Z}_t = dF(q_t,p_t)^* Z_t = \begin{pmatrix} \partial_p(\partial_q H_r)\cdot\tilde{q} - \partial_q(\partial_q H_r)\cdot\tilde{p} \\ \partial_p(\partial_p H_r)\cdot\tilde{q} - \partial_q(\partial_p H_r)\cdot\tilde{p} \end{pmatrix} \\ Z_1 = (\tilde{q}_1, \tilde{p}_1) = (\nabla_{q_1} U, 0) \end{cases} \tag{1.9}$$

we obtain that $\tilde{p}_0 = \nabla_{p_0} U(q_1)$. We can summarize the computation of $\nabla_{p_0} E$ as follows:

Algorithm 1 Compute the gradient $\nabla_{p_0} E$

1: Solve the geodesic equation (1.7) forward in time
2: Compute $\nabla_{q_1} U$
3: Solve system (1.9) backward in time
4: Obtain $\nabla_{p_0} E = \xi_{q_0}\mathbf{K}_V\xi_{q_0}^* p_0 + \tilde{p}_0$.

1.2.3 From landmark to geometric measure matching

As we discussed in the previous sections, the inexact registration of LDDMM relies on two key features: the *diffeomorphic action* and the fidelity term. The action specifies how shapes are deformed by the diffeomorphisms and the fidelity term drives the matching process by penalizing the discrepancy between deformed and the target shapes. For an arbitrary class of shapes to be analyzed, if proper action and fidelity are defined, we can apply the LDDMM model two compare shapes within this type of data. In this section, we briefly review several types of classical examples of shape registration with LDDMM, which are closely related to this thesis.

Landmark points matching. [46] The landmarks are labeled point set so that the correspondence between two sets of landmarks are known. Given the source and target landmarks $x = (x_1, \cdots, x_d)$ and $y = (y_1, \cdots, y_d)$ in $M = \mathbb{R}^{nd}$, the action of the diffeomorphisms on the landmarks is defined as discussed in example 1, so the evolution of the landmarks is given by

$$\begin{aligned} x(t) &= (x_1(t), \cdots, x_d(t)) \\ &\doteq \varphi_t^v \cdot x = (\varphi_t^v(x_1), \cdots, \varphi_t^v(x_d)) \end{aligned}$$

and the data attachment term can be simply defined as the squared Euclidean distance:

$$A(x(1), y) \doteq \lambda \sum_{i=1}^{d} |x_i(1) - y_i|^2,$$

where λ is a weight parameter between the energy and the fidelity term. The geodesic equation or the Hamiltonian equation for landmark matching is as follows:

$$\begin{cases} \dot{x}_i(t) = v_t(x_i(t)) \\ \dot{p}_i(t) = d_{x_i(t)} v_t^T p_i(t) \\ v_t(\cdot) = \sum_{i=1}^{d} k_V(x_i(t), \cdot) p_i(t). \end{cases} \tag{1.10}$$

Unlabeled points matching. [47] The data attachment term for the landmark matching relies on the correspondence between points. Nevertheless, in practical situations, such correspondence may not exist since point sets may not contain the

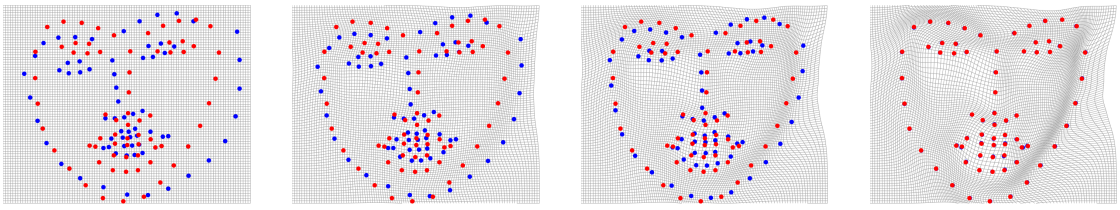


Figure 1-4. An example of LDDMM landmark matching in $\mathbb{R}^2$

same number of points. For example, discrete curves or surfaces data are usually stored as point sets in $2D$ or $3D$ with adjacency matrices and in general there is no correspondence between points. To address this problem, a fidelity term which does not depend on the point to point correspondence is needed. In the machine learning or statistics literature, people often represent a pair of point sets as empirical distribution (as discussed in example 3), and compare them by the maximum mean discrepancy (MMD). See for example [34, 48, 49]. MMD is a type of Hilbert metrics on the space of Borel probability measures obtained from the kernel metric we discussed in section 1.2.1. This type of metrics actually can be defined on the space of Radon measures rather than just probability measures. Let us briefly discuss how to build such Hilbert metrics for Radon measures in $\mathbb{R}^n$. We first choose a kernel $k_W : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ such that $k_W(x, \cdot) \in C_0(\mathbb{R}^n, \mathbb{R}),\ \forall x \in \mathbb{R}^n$. Let W be the RKHS associated with k_W, then from Theorem 5, W is continuously embedded in $C_0(\mathbb{R}^n, \mathbb{R})$. Now, we let $\iota_W : W \hookrightarrow C_0(\mathbb{R}^n, \mathbb{R})$ denote this continuous embedding and $\iota_W^* : C_0(\mathbb{R}^n, \mathbb{R}) \mapsto W^*$ its adjoint, where W^* is the dual space of W. One can compare any two Radon measures μ, μ' through the Hilbert norm of W^* by defining:

$$d_{W^*}(\mu, \mu')^2 = \|\iota_W^* \mu - \iota_W^* \mu'\|_{W^*}^2 \tag{1.11}$$

Note that ι_W^* is not injective in general, which results in that d_{W^*} only induces a pseudo-metric on space of Radon measures. Now consider two unlabelled point sets $X = \{x_1, \cdots, x_N\}$ and $Y = \{y_1, \cdots, y_{N'}\}$, which can be represented by their empirical

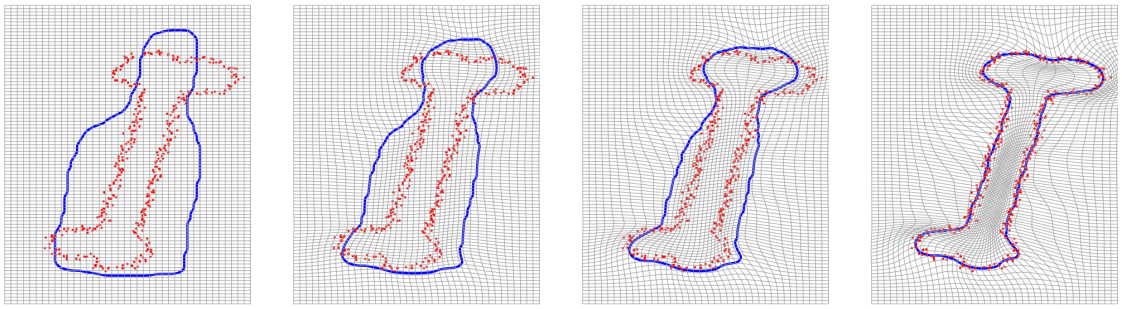
distributions $\mu_X = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$ and $\mu_Y = \frac{1}{N'}\sum_{j=1}^{N'}\delta_{y_j}$. Then the empirical distributions of these two point sets can be compared by this metric:

$$\begin{aligned} &d_{W^*}(\mu_X, \mu_Y) \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} k_W(x_i, x_j) - \frac{2}{NN'}\sum_{i=1}^{N}\sum_{j=1}^{N'} k_W(x_i, y_j) + \frac{1}{(N')^2}\sum_{i=1}^{N'}\sum_{j=1}^{N'} k_W(y_i, y_j). \end{aligned}$$

In [47], the authors adopted the above kernel metric for measures as the fidelity term. Under this framework, the matching problem for unlabelled point sets X and Y is formulated as follows:

$$\begin{aligned} &\min_{v\in L^2([0,1],V)} \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \lambda d_{W^*}(\mu_{X(1)}, \mu_Y) \\ &\text{s.t. } X(t) = \varphi_t^v(X). \end{aligned} \tag{1.12}$$

From the state equation above, the geodesic equation is basically the same as (1.10). The representation of empirical distribution for the point sets assumes uniform weights on the point set. However, if we have further knowledge about the shape of the point sets we can also assign weights to each point. To be concrete, we can represent $X \mapsto \mu_X = \sum_{i=1}^{N} a_i^X \delta_{x_i}$, where $a_i^X > 0,\ \forall i$. For example, in the case of discrete curves or surfaces, we may obtain weights at each point from the length of each segment or area of each triangular mesh. Figure 1-5 illustrates an example of register a curve to a curve like noisy point set.



$t = 0$ $t = 1/3$ $t = 2/3$ $t = 1.$

Figure 1-5. Unlabeled point sets matching in $\mathbb{R}^2$

Submanifold matching. As discussed above, the measure matching approach can be applied to discrete curves or surfaces. Such measure representation and registration can be extended to the continuous case of submanifolds. Let us denote by $\mathcal{H}^d$ the d-dimensional **Hausdorff measure** in $\mathbb{R}^n$. $\mathcal{H}^d$ is a measure on $\mathbb{R}^n$ which basically measures the d-dimensional volume of subsets in $\mathbb{R}^n$. When $d = n$, $\mathcal{H}^n$ is the usual Lebesgue measure and when $d = 0$, $\mathcal{H}^n$ is the *counting measure*. If X is a ℓ-dimensional submanifold in $\mathbb{R}^n$, then

$$\mathcal{H}^d(X) = \begin{cases} d\text{-volume of } X, & \text{if } \ell = d \\ 0, & \text{if } \ell < d \\ \infty, & \text{if } \ell > d. \end{cases}$$

For the precise construction and properties of Hausdorff measures, we refer the reader to [50, 51].

Now for any d-dimensional submanifold X, the measure associated with X is defined as a measure defining as the d-dimensional Hausdorff measures restricted to X:

$$\mu_X(A) \doteq (\mathcal{H}^d) \llcorner X(A) = \mathcal{H}^d(X \cap A).$$

Note that it is well known that μ_X introduced above coincides with the "volume" in X. To be precise, let $\psi : U \subset \mathbb{R}^d \mapsto X$ be a local chart and for any Borel subset $B \subset \psi(U)$, we can define its volume by

$$\mathrm{vol}_X(B) \doteq \int_{\psi^{-1}(B)} \sqrt{\det((d\psi)^* \circ d\psi)} dx.$$

Assuming $X \cap A \subset \psi(U)$ as well, then $\mu_X(A) = \mathrm{vol}_X(A \cap X)$. With the above representation, the registration problem for d-dimensional submanifolds X and Y is then formulated as (1.12).

In summary, the central idea of this approach is to make use of the embedding of unparametrized shapes into the space of measures and build **parametrization-invariant** fidelity metrics as restrictions of metrics on those measure spaces themselves.

However, such measure representation of shapes only regard submanifolds as subsets of $\mathbb{R}^n$ and only take into account the distribution of points in the shapes. The local and global volume changes of the shapes made by the deformation during the morphing process are not captured by such measure embedding.

To address this issue, several competing approaches have been introduced, each relying on embeddings into different spaces of generalized measures: [4, 52, 53] are based on the representation of oriented curves and surfaces as *currents*, [52, 53] and [45, 54–56] extended this model to the setting of unoriented and oriented *varifolds*, while [57–59] considers the higher-order representation of *normal cycles*, see also the recent survey [60]. One common feature of all those works, however, is that they are focused primarily on registrations of curves or surfaces. In other words, the use of current, varifolds, or normal cycles confines to the computation of a fidelity metric to guide registration algorithms but the deformation model itself remains tied to the curve/surface setting or equivalently, in the discrete situation, to objects described by point set meshes.

The guiding theme and main objective of this thesis are to investigate an alternative framework that, in contrast with those prior works, would formulate the deformation model as well as tackle the registration problem directly in these generalized measure spaces: we focus specifically on the (oriented) varifold setting of [45, 55]. We shall recap the main definition and properties related to varifolds in the following section.

1.3 An overview of varifolds in shape analysis

The concept of varifold was originally developed in the context of geometric measure theory by [61], [62] and [63] for the study of Plateau's problem on minimal surfaces. The interest in registration and shape analysis was evidenced in [45, 54, 55]. In those works, varifolds provide a convenient representation of geometric shapes such

as rectifiable curves and surfaces and an efficient approach to define and compute fidelity terms for registration, or to perform clustering, classification in those shape spaces. The main purpose of this section is to introduce varifolds in this latter context. The case of non-oriented shapes was thoroughly investigated in [45]. Later on, the generalized framework of oriented varifold was proposed in [55] but only for objects of dimension or co-dimension one. In the following, we provide a fully general presentation of oriented varifolds and their properties, that also does not specifically focus on the case of rectifiable varifolds as these previous works did. Although we assume here that all the considered shapes are oriented, we emphasize that the non-oriented framework of [45] can be recovered almost straightforwardly through adequate choices of orientation-invariant kernels as we shall briefly point out later on.

1.3.1 Varifold representation of shapes

The underlying principle of varifolds is to extend measures of $\mathbb{R}^n$ by incorporating an additional tangent space component. In this thesis, we will consider such spaces to be oriented. Thus, for a given dimension $0 \leq d \leq n$, we first need to introduce the set of all possible d-dimensional oriented tangent spaces in $\mathbb{R}^n$:

Definition 4. *The d-dimensional oriented Grassmannian $\widetilde{G}_d^n$ is the set of all oriented d-dimensional linear subspaces of $\mathbb{R}^n$.*

Remark 1. *The oriented Grassmannian is a compact manifold of dimension $d(n-d)$ which can be identified to the quotient $SO(n)/(SO(d) \times SO(n-d))$. It is also a double cover of the (non-oriented) Grassmannian G_d^n of d-dimensional subspaces of $\mathbb{R}^n$.*

For practical purposes, we need a more convenient representation of $\widetilde{G}_d^n$ by the *d-vectors*. This requires some knowledge from multilinear algebra and we shall briefly recap some properties on this topic following Chapter 4 in [64]. More detailed discussions can also be found in [65]. Let $u_1, \cdots, u_d$ be vectors in $\mathbb{R}^n$. One can define

a way to multiply these m-vectors to obtain a new object called a *d-vector*:

$$u_1 \wedge \cdots \wedge u_d.$$

This wedge product has two important properties, multilinearity:

$$cu_1 \wedge u_2 = u_1 \wedge cu_2 = c(u_1 \wedge u_2)$$
$$(u_1 + u_2) \wedge (v_1 + v_2) = u_1 \wedge v_1 + u_1 \wedge v_2 + u_2 \wedge v_1 + u_2 \wedge v_2.$$

and alternating:

$$u \wedge v = -v \wedge u.$$

Let $e_1, \cdots, e_n$ be the standard basis of $\mathbb{R}^n$. From the properties mentioned above, one can verify that

$$u_1 \wedge \cdots \wedge u_d = \sum_{i_1 < \cdots < i_d} a_{i_1 \cdots i_d} e_{i_1} \wedge \cdots \wedge e_{i_d}.$$

The set of all linear combination of $\{e_{i_1} \wedge \cdots \wedge e_{i_d}\}_{i_1 < \cdots < i_d}$ is denoted by $\Lambda^d(\mathbb{R}^n)$ and the elements in $\Lambda^d(\mathbb{R}^n)$ are called *d-vectors*. A natural inner product can be defined on $\Lambda^d(\mathbb{R}^n)$ so that $\{e_{i_1} \wedge \cdots \wedge e_{i_d}\}_{i_1 < \cdots < i_d}$ is an orthonormal basis. It can be showed that, the inner product between any $\xi = \xi_1 \wedge \ldots \wedge \xi_d$, $\eta = \eta_1 \wedge \ldots \wedge \eta_d$ in $\Lambda^d(\mathbb{R}^n)$ can be expressed as the determinant of the Gram matrix:

$$\langle \xi, \eta \rangle = \det(\xi_i \cdot \eta_j)_{i,j=1,\ldots,d}. \tag{1.13}$$

Given $T \in \widetilde{G}_d^n$, there exists a basis $\{u_i\}_{i=1,\ldots,d} \in \mathbb{R}^{n \times d}$ of T such that $\{u_1, \cdots, u_d\}$ has consistent orientation with T. Then the following map, called the oriented Plücker embedding, is well defined and injective,

$$i_P : \widetilde{G}_d^n \mapsto \{\xi \in \Lambda^d(\mathbb{R}^n) : |\xi| = 1\}$$
$$T \mapsto \frac{u_1 \wedge \cdots \wedge u_d}{|u_1 \wedge \cdots \wedge u_d|}.$$

This allows to identify $\widetilde{G}_d^n$ as a subset of the unit sphere of $\Lambda^d(\mathbb{R}^n)$ which inherits the topology of the inner product on $\Lambda^d(\mathbb{R}^n)$. Through this identification, one can also define the action of linear transformations on $\widetilde{G}_d^n$ as follows

$$A \cdot T := \frac{Au_1 \wedge \cdots \wedge Au_d}{|Au_1 \wedge \cdots \wedge Au_d|} \tag{1.14}$$

for any $T \in \widetilde{G}_d^n$ and $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ a linear invertible map.

Definition 5. *An oriented d-varifold μ on $\mathbb{R}^n$ is a nonnegative finite Radon measure on the space $\mathbb{R}^n \times \widetilde{G}_d^n$. Its weight measure $|\mu|$ is defined by $|\mu|(A) := \mu(A \times \widetilde{G}_d^n)$ for all Borel subset A of $\mathbb{R}^n$. We denote by $\mathcal{V}_d$ the space of all oriented d-varifolds.*

In the rest of this thesis, with a slight abuse of vocabulary, we will often use the word varifold instead of oriented varifold for the sake of concision. Recall that from the Riesz representation theorem, we can alternatively view any varifold μ as a distribution, i.e. an element of the dual space $C_0(\mathbb{R}^d \times \widetilde{G}_d^n)^*$, where $C_0(\mathbb{R}^d \times \widetilde{G}_d^n)$ denotes the set of continuous functions vanishing at infinity on $\mathbb{R}^d \times \widetilde{G}_d^n$. It is defined for any test function $\omega \in C_0(\mathbb{R}^d \times \widetilde{G}_d^n)$ by:

$$(\mu|\omega) \doteq \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \omega(x,T) d\mu(x,T). \tag{1.15}$$

As an additional note, another useful representation of a general varifold in $\mathcal{V}_d$ can be obtained by the disintegration theorem (see [66] Chap. 2). Namely, if $\mu \in \mathcal{V}_d$, for $|\mu|$-almost every x in $\mathbb{R}^n$, there exists a probability measure ν_x on $\widetilde{G}_d^n$ such that $x \mapsto \nu_x$ is $|\mu|$-measurable and we can write

$$(\mu|\omega) = \int_{\mathbb{R}^n} \int_{\widetilde{G}_d^n} \omega(x,T) d\nu_x(T) d|\mu|(x). \tag{1.16}$$

In other words, the varifold μ can be decomposed as its weight measure on $\mathbb{R}^n$ together with a family of tangent space probability measures on the Grassmannian at the different points in the support of $|\mu|$. This is usually referred to as the Young measure representation of μ.

There are a few important subfamilies of varifolds which will be relevant for the following. The first particular class is the one of *rectifiable varifolds*, which are in essence the varifolds representing an oriented shape of dimension d. More precisely, given an oriented d-dimensional submanifold X of $\mathbb{R}^n$ of finite total d-volume, denoting by $T_X(x) \in \widetilde{G}_d^n$ the oriented tangent space at $x \in X$, one can associate to X the varifold μ_X, which is defined for all Borel subset $B \subset \mathbb{R}^n \times \widetilde{G}_d^n$ by $\mu_X(B) = \mathcal{H}^d(\{x \in X | (x, T_X(x)) \in B\})$. It is then not hard to see that, as an element of $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)^*$,

$$\begin{aligned}(\mu_X|\omega) &= \int_{\mathbb{R}^d \times \widetilde{G}_d^n} \omega(x,T) d\mu_X(x,T) \\ &= \int_X \omega(x, T_X(x)) d\mathcal{H}^d(x). \end{aligned} \tag{1.17}$$

Such a representation $X \mapsto \mu_X$ can be extended to slightly more general objects known as *oriented rectifiable sets*.

Definition 6. *A subset X of $\mathbb{R}^n$ is said to be a countably d-rectifiable set if $\mathcal{H}^d(X \setminus \cup_{j=1}^{\infty} F_j(\mathbb{R}^d)) = 0$, where $F_j : \mathbb{R}^d \mapsto \mathbb{R}^n$ are Lipschitz function for all j (cf. [50]). We say that (X, T_X) is an oriented rectifiable set if X is a countably d-rectifiable set and $T_X : X \mapsto \widetilde{G}_d^n$ is a $\mathcal{H}^d$-measurable function such that for $\mathcal{H}^d - a.e.$ $x \in X$, $T_X(x)$ is the approximate tangent space of X at x with specified orientation.*

Rectifiable subsets include both usual submanifolds but also piecewise smooth objects like polyhedra. Given any oriented rectifiable set (X, T_X), we can associate a varifold that we also write μ_X given again by (1.17). The set of those μ_X will be referred to as the rectifiable oriented varifolds in this thesis (note that this is actually more restrictive than the standard definition of rectifiable varifold in the literature which also incorporates an additional multiplicity function).

Remark 2. *Rectifiable varifolds still make up for a very "small" subset of $\mathcal{V}_d$: indeed, in the Young measure representation of (1.16), we have in this case the very particular constraint that probability measures ν_x are Dirac masses, specifically $\nu_x = \delta_{T_X(x)}$.*

Second of those are the Diracs. For $x \in \mathbb{R}^n$ and $T \in \widetilde{G}_d^n$, the Dirac varifold $\delta_{(x,T)}$ acts on functions of $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$ by the relation

$$(\delta_{(x,T)}|\omega) = \omega(x,T),\ \forall \omega \in C_0(\mathbb{R}^n \times \widetilde{G}_d^n).$$

$\delta_{(x,T)}$ can be viewed as a singular particle of mass 1 at position x that carries the oriented d-plane T. We will also consider a specific class of varifolds which can be written as finite combinations of Dirac masses:

Definition 7. *A varifold μ is called a discrete varifold if it has the following form:*

$$\mu = \sum_{i=1}^{N} r_i \delta_{(x_i,T_i)},\ r_i \in \mathbb{R}_+,\ x_i \in \mathbb{R}^n,\ T_i \in \tilde{G}_d^n, \tag{1.18}$$

for some $N \geq 1$.

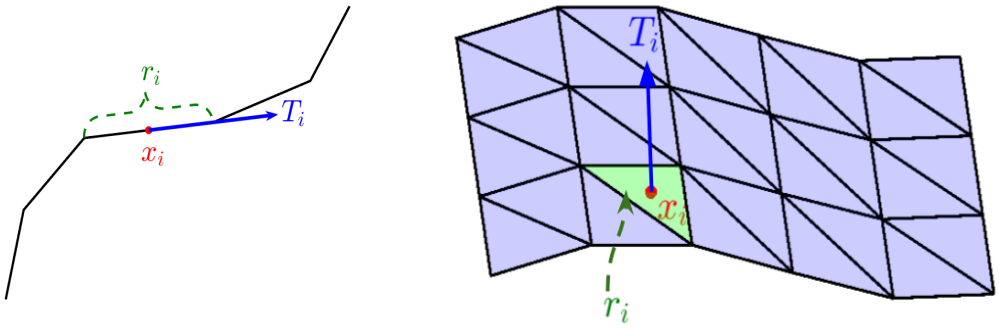


Figure 1-6. Discrete varifold representation of discrete curves and surfaces.

It is quite natural to consider this type of varifolds for the purpose of representing discrete shapes, which has been exploited in previous works on piecewise linear curves and surfaces. For example, if $X = \bigcup_{i=1}^{N} X_i$ is a triangulated surface, with X_i being the mesh triangles with specified orientations, one can write $\mu_X = \sum_{i=1}^{N} \mu_{X_i}$ and for each $i \in \{1, \cdots, N\}$ approximate μ_{X_i} by $r_i \delta_{(x_i,T_i)}$, where x_i is the center of X_i, T_i the oriented plane containing X_i and $r_i = \mathcal{H}^d(X_i)$. This leads to the approximation $\widetilde{\mu}_X := \sum_{i=1}^{N} r_i \delta_{(x_i,d_i)}$ as illustrated in figure 1-6. As proved in [55], this approximation

provides an acceptable error bound for the kernel metric d_{W^*}, which will be introduced in section 1.3.3:

$$d_{W^*}(\mu_X, \widetilde{\mu}_X) \leq Cte\ \mathcal{H}^d(X) \max_i \operatorname{diam}(X_i).$$

1.3.2 Transportation of varifolds by diffeomorphisms

In this section, we discuss different models for how varifolds can be transported by a diffeomorphism of $\mathbb{R}^n$, in other words what are possible group actions of the diffeomorphism group $\mathrm{Diff}(\mathbb{R}^n)$ on $\mathcal{V}_d$.

Let us start by considering the case of an oriented rectifiable subset (X, T_X). A diffeomorphism $\phi \in \mathrm{Diff}(\mathbb{R}^n)$ transports (X, T_X) as

$$\phi \cdot (X, T_X) \doteq (\phi(X), T_{\phi(X)}),$$

where the transported orientation map writes

$$T_{\phi(X)}(y) \doteq d_{\phi^{-1}(y)}\phi \cdot T_X(\phi^{-1}(y))$$

the above term being well-defined from (1.14). This suggests introducing the following *pushforward action* on $\mathcal{V}_d$, which is defined for all $\mu \in \mathcal{V}_d$ and $\phi \in \mathrm{Diff}(\mathbb{R}^n)$ by:

$$(\phi_\# \mu | \omega) \doteq \int_{\mathbb{R}^d \times \widetilde{G}_d^n} \omega(\phi(x), d_x\phi \cdot T) J_T\phi(x) d\mu(x, T) \tag{1.19}$$

in which $J_T\phi(x)$ denotes the determinant of the Jacobian of ϕ along T (i.e. the change of d-volume induced by ϕ along T at x) which is given by

$$J_T\phi(x) = \sqrt{\det\left((d_x\phi(e_i) \cdot d_x\phi(e_j))_{i,j=1,\dots,d}\right)}$$

for $(e_1, \dots, e_d)$ an orthonormal basis of T. One easily verifies that $(\phi, \mu) \mapsto \phi_\#\mu$ defines a group action which commutes with the action on oriented rectifiable sets, namely

Proposition 2. *For any oriented rectifiable set (X, T_X) and diffeomorphism $\phi \in$ Diff$(\mathbb{R}^n)$, $\phi_\#\mu_X = \mu_{\phi(X)}$.*

This follows from the following **area formula** for integrals over rectifiable sets, c.f. [50] Chapter 2 or [67] Chapter 3. We provide the proof of Proposition 2 in the appendix.

Theorem 8. *Let X be a countably r-rectifiable set contained in an open set U in $\mathbb{R}^\ell$ and $f : U \mapsto \mathbb{R}^m$ be a locally Lipschitz map with $m \geq d$. Then for any $\mathcal{H}^r$-measurable function h, we have the following area formula:*

$$\int_X h(x) J_{T_x X} f(x) d\mathcal{H}^r(x) = \int_{f(X)} \int_{f^{-1}(y)} h(x) d\mathcal{H}^0(x) d\mathcal{H}^r(y). \tag{1.20}$$

It also straightforward to see that the transport of a Dirac mass $\delta_{(x,T)}$ is given by:

$$\phi_\# \delta_{(x,T)} = |J_T \phi(x)| \delta_{(\phi(x), d_x \phi \cdot T)}. \tag{1.21}$$

Figure 1-7 is an illustration of transporting a Dirac mass by a flow of diffeomorphisms through this pushforward action.

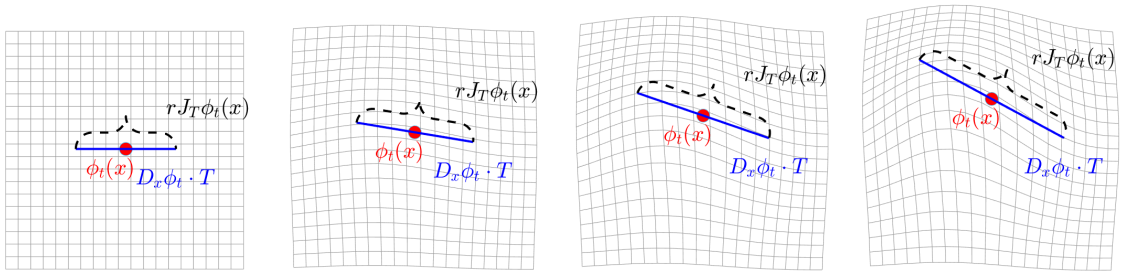


Figure 1-7. A Dirac mass transported by a flow of diffeomorphisms through the pushforward action.

Remark 3. *This pushforward action also extends the diffeomorphic transport of measures with densities on $\mathbb{R}^n$. Indeed if $\mu = \theta(x).\mathcal{L}^n$ with θ a measurable density function on $\mathbb{R}^n$ and $\mathcal{L}^n$ the Lebesgue measure, we can extend μ to a 0-varifold in $\mathcal{V}_0$ by taking a constant global orientation in $\widetilde{G}_0^n = \{\pm 1\}$, say $+1$ for all $x \in \mathbb{R}^n$. Then, for any orientation-preserving diffeomorphism ϕ, (1.19) in this case can be simplified as:*

$$(\phi_\# \mu | \omega) = \int_{\mathbb{R}^d} \omega(\phi(x)) \theta(x) dx = \int_{\mathbb{R}^d} \omega(y) J\phi^{-1}(y) \theta \circ \phi^{-1}(y) dy$$

where the second equality follows from the change of variable $y = \phi(x)$. Therefore we see that $\phi_{\#}\mu$ is also density measure and the density is $J\phi^{-1}(x)\theta \circ \phi^{-1}(x)$ with $J\phi^{-1}(x)$ being the full Jacobian determinant of ϕ^{-1}. This is consistent with the usual diffeomorphic transformation of densities discussed for instance in [21] Chapter 9.6.1.

1.3.3 Metrics on varifolds

In this section, we address the issue of defining adequate metrics on the space $\mathcal{V}_d$. After reviewing some classical metrics and their limitations for the specific applications of this work, we turn to metrics defined through positive definite kernels, for which we extend previous constructions introduced in e.g. [45, 55]. Further properties of this class of distances and its relationship to the classical metrics will be explored in Chapter 3.

As a measure/distribution space, $\mathcal{V}_d$ can be equipped with various topologies and metrics, several of which have been regularly used in various contexts. We discuss a few of those below.

- *mass norm*: with the previous identification of measures in $\mathcal{V}_d$ with elements of the dual $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)^*$, one can define the following dual metric on $\mathcal{V}_d$:

$$d_{op}(\mu, \nu) \doteq \sup_{|\omega|_\infty \leq 1} (\mu - \nu | \omega), \ \forall \mu \in \mathcal{V}_d. \tag{1.22}$$

  where $|\omega|_\infty \doteq \sup_{\mathbb{R}^n \times \widetilde{G}_d^n} |\omega|$. This metric is generally too strong for applications in shape analysis and leads to a discontinuous behavior. Indeed, one can easily verify that for any two Dirac masses $\delta_{(x,T)}$ and $\delta_{(x',T')}$, $d_{op}(\delta_{(x,T)}, \delta_{(x',T')}) = 2$ whenever $(x, T) \neq (x', T')$.

- *weak-** topology: a sequence of d-varifolds $\{\mu_i\}_i$ converges to $\mu \in \mathcal{V}_d$ in the weak-* topology (denoted by $\mu_i \overset{*}{\rightharpoonup} \mu$) if and only if for all $\omega \in C_c(\mathbb{R}^d \times \widetilde{G}_d^n)$

(continuous compactly supported function)

$$\lim_{i\to\infty}(\mu_i|\omega) = (\mu|\omega). \tag{1.23}$$

In fact, the weak-* topology on $\mathcal{V}_d$ can be metrized by the following distance:

$$d_*(\mu,\nu) = \sum_{k\in\mathbb{N}} 2^{-k}|(\mu-\nu|\omega_k)|,$$

where $\{\omega_k\}_{k\in\mathbb{N}}$ is a dense sequence in $C_c(\mathbb{R}^n \times \widetilde{G}_d^n)$.

- *Wasserstein metric*: the Wasserstein-1 distance of optimal transport can be expressed in its Kantorovitch dual formulation [68] as

$$d_{Wass^1}(\mu,\nu) \doteq \sup_{\mathrm{Lip}(\omega)\leq 1} |(\mu-\nu|\omega)|. \tag{1.24}$$

where the sup is taken over all Lipschitz regular functions on $\mathbb{R}^n \times \widetilde{G}_d^n$ with Lipschitz constant smaller than one. This metric is however well-suited for measures with the same total mass. Several recent works [69, 70] have instead proposed generalized Wasserstein distances derived from unbalanced optimal transport.

- *Bounded Lipschitz metric*: similar to the previous, the bounded Lipschitz distance (sometimes referred to as the *flat metric*) on $\mathcal{V}_d$ is defined by

$$d_{BL}(\mu,\nu) \doteq \sup_{\|\omega\|_\infty,\mathrm{Lip}(\omega)\leq 1} |(\mu-\nu|\omega)|. \tag{1.25}$$

It can be shown (cf. Ch 8 in [71]) that d_{BL} metrizes the *narrow topology* on $\mathcal{V}_d$, namely the topology for which a sequence (μ_i) converges to μ if and only if $\lim_{i\to\infty}(\mu_i|\omega) = (\mu|\omega)$ for all bounded continuous functions ω.

Clearly, the narrow topology is stronger than the weak-* topology. Furthermore, it is also well known that d_{BL} locally metrizes the weak-* topology on $\mathcal{V}_d$, namely:

Proposition 3. *Let μ and $\{\mu_i\}_i$ be varifolds such that the sequence $\{\mu_i\}_i$ is tight. Then $\mu_i \overset{*}{\rightharpoonup} \mu$ if and only if $d_{BL}(\mu_i,\mu)\to 0$.*

As a direct consequence of Proposition 3, we have in particular that weak-* convergence and convergence in d_{BL} are equivalent if one restricts to varifolds that are supported in a fixed compact subset of $\mathbb{R}^n \times \widetilde{G}_d^n$. Note also that a very similar result to Proposition 3 holds when replacing the bounded Lipschitz distance by generalized Wasserstein metrics, as proved in [69].

The above metrics on varifolds all originate from classical ones in standard measure theory. Unlike the mass norm, Wasserstein and bounded Lipschitz metrics have nice theoretical properties in terms of shape comparison. However, for the purpose of diffeomorphic registration that we shall tackle below, one needs metrics that are easy to evaluate numerically. This is typically not the case of d_{Wass^1} and d_{BL} expressed above as there is no straightforward way to compute the corresponding suprema over the respective sets of test functions. One line of work has been considering approximations of optimal transport distances with e.g. entropic regularizers for which Sinkhorn-based algorithms can be derived, see for instance the recent work [72] as well as the recent survey [73] of computational frameworks for optimal transport.

In this thesis, we focus on the alternative approach previously developed for currents in [53] and unoriented varifolds in [45] which instead defines a general class of pseudo-metrics on $\mathcal{V}_d$ based on positive definite kernels and their corresponding *reproducing kernel Hilbert space* (RKHS). In the context of varifolds, we are interested in defining positive definite kernels on the product $\mathbb{R}^n \times \widetilde{G}_d^n$. Along the lines of previous works like [45, 55], we build kernels from the tensor product of kernels on $\mathbb{R}^n$ and $\widetilde{G}_d^n$. To be concrete, let k^{pos} and k^G be positive definite kernels on $\mathbb{R}^n$ and $\widetilde{G}_d^n$ respectively, it is straightforward to verify that the tensor product $k \doteq k^{pos} \otimes k^G$ defined as follows,

$$k((x,T),(x',T')) \doteq k^{pos}(x,x')k^G(T,T')$$

is a positive kernel on $\mathbb{R}^n \times \widetilde{G}_d^n$. The construction of kernels on the Grassmannian manifolds has been extensively studied in the literature, see for example [74–76]. With

some continuity assumptions on the kernels k^{pos} and k^G, one can further show that k is continuously embedded in $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$:

Proposition 4. *Let k^{pos} and k^G be continuous positive definite kernels on $\mathbb{R}^n$ and $\widetilde{G}_d^n$ respectively. Assume in addition that for any $x \in \mathbb{R}^n$, $k^{pos}(x, \cdot) \in C_0(\mathbb{R}^n)$. Then $k := k^{pos} \otimes k^G$ is a positive definite kernel on $\mathbb{R}^n \times \widetilde{G}_d^n$ and the RKHS W associated to k is continuously embedded in $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$ i.e. there exists $c_W > 0$ such that for any $\omega \in W$, we have $\|\omega\|_\infty \leq c_W \|\omega\|_W$.*

The proof of Proposition 4 can be found in the appendix. From now on, we assume that the kernels on $\mathbb{R} \times \widetilde{G}_d^n$ satisfy the assumptions in Proposition 4.

Now, if we let $\iota_W : W \hookrightarrow C_0(\mathbb{R}^d \times \widetilde{G}_d^n)$ be the continuous embedding given by Proposition 4 and ι_W^* be its adjoint, then for any $\mu \in C_0(\mathbb{R}^n \times \widetilde{G}_d^n)^*$, we have

$$(\iota_W^* \mu | \omega) = \int_{\mathbb{R}^d \times \widetilde{G}_d^n} \omega(x, T) d\mu(x, T), \ \forall \omega \in W. \tag{1.26}$$

With (1.26), we may identify μ as an element of the dual RKHS W^*. Note that ι_W^* is not injective in general, in other words one can have $\mu = \mu'$ in W^* but $\mu \neq \mu'$ in $C_0(\mathbb{R}^n \times \widetilde{G}_d^n)^*$.

In any case, one can compare any two varifolds $\mu, \mu' \in \mathcal{V}_d$ through the Hilbert norm of W^* by defining:

$$d_{W^*}(\mu, \mu')^2 = \|\mu - \mu'\|_{W^*}^2 = \|\mu\|_{W^*}^2 - 2\langle \mu, \mu' \rangle_{W^*} + \|\mu'\|_{W^*}^2 \tag{1.27}$$

where we use the small abuse of notation of writing μ and μ' instead of $\iota_W^* \mu$ and $\iota_W^* \mu'$ on the two right hand sides. Due to the potential non-injectivity of ι_W^*, in general d_{W^*} only induces a pseudo-metric on $\mathcal{V}_d$.

To simplify the rest of the presentation and in the perspective of later numerical considerations, we will also assume specific forms for k^{pos} and k^G, namely that k^{pos} is a translation/rotation invariant radial kernel $k^{pos}(x, y) = \rho(|x - y|^2), \ \forall x, y \in \mathbb{R}^n$,

with $\rho(0) > 0$, and k^G is a zonal kernel $k^G(S,T) = \gamma(\langle S,T\rangle),\ \forall S,T \in \widetilde{G}_d^n$ where $\langle\cdot,\cdot\rangle$ is the inner product on $\widetilde{G}_d^n$ inherited from $\Lambda^d(\mathbb{R}^n)$ introduced in section 1.3.1. In this thesis, we will refer to this type of kernels as **R-Z kernel** and we summarize the above assumptions in the following definition:

Definition 8. *A kernel $k : (\mathbb{R}^n \times \widetilde{G}_d^n) \mapsto \mathbb{R}$ is called a **R-Z kernel** if it has the following form,*

$$k((x,T),(x',T')) = k^{pos}(x,x')k^G(\langle T,T'\rangle) = \rho(|x-x'|^2)\gamma(\langle T,T'\rangle), \tag{1.28}$$

with $\rho(0) > 0$.

These assumptions are quite natural as they will eventually induce metrics on varifolds invariant to the action of rigid motion. Note that the unoriented framework of [45] can be also recovered in this setting by simply restricting to orientation-invariant kernels k^G i.e. such that $\gamma(-t) = \gamma(t)$ for all t.

The main advantage of this construction is that d_{W^*} can be now expressed explicitly based on the reproducing kernel property of W. Indeed, given any μ and ν in $\mathcal{V}_d$, the inner product between them is given by

$$\begin{aligned}\langle \mu,\mu'\rangle_{W^*} &= \int_{(\mathbb{R}^d\times\widetilde{G}_d^n)^2} k^{pos}(x,x')k^G(T,T')d\mu(x,T)d\mu'(x',T') \\ &= \int_{(\mathbb{R}^d\times\widetilde{G}_d^n)^2} \rho(|x-x'|^2)\gamma(\langle T,T'\rangle)d\mu(x,T)d\mu'(x',T')\end{aligned} \tag{1.29}$$

for kernels selected as discussed in the paragraph above. When μ and μ' are discrete varifolds, the metric (1.29) becomes particularly simple to compute numerically. Indeed, if $\mu = \sum_{i=1}^N r_i\delta_{(x_i,S_i)}$ and $\mu' = \sum_{j=1}^M r'_j\delta_{(x'_j,T'_j)}$, we have as a particular case of (1.29):

$$\langle \mu,\mu'\rangle_{W^*} = \sum_{i=1}^N\sum_{j=1}^M r_i r'_j \rho(|x_i-x'_j|^2)\gamma(\langle T_i,T'_j\rangle). \tag{1.30}$$

Going back to the submanifold registration problem introduced at the end of the previous section, kernel metrics on varifolds provide a convenient class of data

attachment terms to measure the discrepancy e.g. between two curves or between two surfaces. In general, given two rectifiable subsets X_0 and X_{tar}, the LDDMM framework for registering X_0 to X_{tar} is very similar to the measure matching approach discussed in section 1.2.3. Thanks to the commutative property stated in Proposition 2, it can be formulated as the following optimal control problem:

$$\underset{v\in L^2([0,1],V)}{\operatorname{argmin}} \left\{ \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \lambda \|\mu_{X(1)} - \mu_{X_{tar}}\|_{W^*}^2 \right\} \tag{1.31}$$

with the state equation $X(t) = \varphi_t^v \cdot X_0$. This is the general setting of several past works in diffeomorphic shape registration including [45, 52, 53, 55].

1.4 Contributions and organization

As we discussed in sections 1.2 and 1.3, the previous works only used the varifold representation as surrogates to define and evaluate the fidelity term for registration, which confines their framework only to the data in the form of curves or surfaces. Unlike the previous approaches, the main objective of this thesis is to develop the general mathematical model and numerical methods for registrations in the space of varifolds. There are several arguments for the interest of such an approach but in our point of view, the primary motivation lies in the fact that varifolds being more general than submanifolds, the proposed framework allows to extend the large deformation analysis methods to a range of new geometric objects while giving more flexibility to deal with some of the flaws which are commonplace in shapes segmented from raw data. We then take advantage of such flexibility to extend our model to handle more general situations such as varifold compression and imbalanced shape matching. The organization of this thesis is as follows:

- **Chapter 2 :** In this chapter, we derive a formulation of LDDMM registration of general varifolds, for which we show the existence of solutions and derive the Hamiltonian equations associated with the corresponding optimal

control problem. In the perspective of numerical implementations, discretization frameworks are needed since, in practice, shapes are represented by discrete varifolds. We derive the discrete version of the optimal control problem and optimality equations, from which we deduce a geodesic shooting algorithm for the diffeomorphic registration of discrete varifolds.

- **Chapter 3 :** First, we propose a comprehensive study of the class of kernel metrics on varifold spaces initiated in [45, 55], in particular by examining the required conditions to recover true distances between all varifolds (as opposed to the subset of rectifiable varifolds) and comparing the resulting topologies with some standard metrics on measures. Next, we address the issue of quantization in varifold space, namely of approximating any varifold as a finite sum of Dirac masses. We consider a novel approach for varifold compression in this context, that consists in computing projections onto particular cones of discrete varifolds. We then make use of the approach developed in chapter 2 to register compressed varifolds and prove the Γ-convergence of the corresponding approximate registration functionals.

- **Chapter 4 :** We extend the approach developed in chapter 1 to more general registration models which incorporate additional density changes in order to tackle registrations of imbalanced shapes. We first derive a global density change model by introducing a rescaling factor which account for the potential global mass imbalance between shapes. Such a model is well-suited when a common and global density rescaling effect is expected but is typically not adapted to the situation of local mass imbalances such as in the case of particular missing parts on the target shape. To tackle this more general case, we first develop a preliminary local density change model which is a simple extension of the previous one. However, this simple approach makes it difficult to define shape

spaces and metrics due to its lack of symmetry. We then propose an alternative model in line with Grenander's shape theory, that defines the weight rescaling through a dynamical process and introduces shape metrics from group actions on varifolds.

Chapter 2

Diffeomorphic registration of varifolds

The goal of this chapter is to generalize the submanifold registration framework presented in section 1.3 and formulated as (1.31). This approach only make use of the varifold representation for defining the data attachment term and is restricted to data with submanifold structure. In this chapter, we derive deformation models for general varifolds in order to formulate and study the diffeomorphic registration problem directly in the space of varifolds on the theoretical level. On the numerical side, we study in detail the optimality conditions for the discrete matching problem and develop numerical methods for registration of discrete varifolds, which allow us to register more general geometric objects that can be represented as distributions of points with directions. Results of this chapter were partially published in [77, 78].

2.1 General framework

2.1.1 The diffeomorphic registration problem

With the group action defined in (1.19), we are now ready to introduce the mathematical formulation of the diffeomorphic registration problem for general varifolds in $\mathcal{V}_d$. As deformation model, we will again rely on the Large Deformation Diffeomorphic Metric Mapping (LDDMM) setting presented in chapter 1.

Consider now a source (or template) varifold $\mu_0 \in \mathcal{V}_d$ as well as a target $\mu_{tar} \in \mathcal{V}_d$. The LDDMM framework of registering μ_0 to μ_{tar} consists in finding a deformation ϕ that minimizes $d_{G_V}(\mathrm{id}, \phi)$ with the constraint that $\phi_{\#}\mu_0$ is close to μ_1 in the sense of a kernel metric $\|\cdot\|_{W^*}$ defined in Section 1.3.3. This can be reformulated as the following optimal control problem:

$$\underset{v \in L^2([0,1],V)}{\operatorname{argmin}} \left\{ E(v) = \frac{1}{2} \int_0^1 \|v_t\|_V^2 dt + \lambda \|\mu(1) - \mu_{tar}\|_{W^*}^2 \right\} \tag{2.1}$$

with v being the control, E the total cost and the state equation is given by $\mu(t) \doteq (\phi_t^v)_{\#}\mu_0$ for the pushforward model. The first term in (2.1) is the regularization term that constrains the regularity of the estimated deformation paths. The second term measures the similarity between the deformed varifold $\mu(1)$ and the target varifold μ_{tar}. λ is a weight parameter between the regularization and fidelity terms.

The well-posedness of the optimal control problem (2.1) holds under the following assumptions:

Theorem 9. *If V is continuously embedded in $C_0^2(\mathbb{R}^n, \mathbb{R}^n)$, W is continuously embedded in $C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ and $\mathrm{supp}(\mu_0) \subset \mathrm{K}$, for some compact subset K of $\mathbb{R}^n \times \widetilde{G}_d^n$, then there exists a global minimizer to the problem (2.1).*

Proof. Thanks to the first term in E, any minimizing sequence of E is bounded in $L^2([0,1], V)$. Let (v^j) be a subsequence of such minimizing sequence which converges weakly to some $\bar{v}$ in $L^2([0,1], V)$. Using the results of [21] Chapter 7.2, we know that

$$\lim_{j \to \infty} \|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} = 0.$$

Furthermore, for any $\omega \in W$, we have

$$\begin{aligned}
&\left|\left((\varphi_1^{v^j})_{\#}\mu_0 - (\varphi_1^{\bar{v}})_{\#}\mu_0|\omega\right)\right| \\
&= \left| \int_K J_S\varphi_1^{v^j}(x)\omega(\varphi_1^{v^j}(x), d_x\varphi_1^{v^j}\cdot S) - J_S\varphi_1^{\bar{v}}(x)\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)d\mu_0 \right| \\
&\leq \int_K |J_S\varphi_1^{v^j}(x)| \left|\omega(\varphi_1^{v^j}(x), d_x\varphi_1^{v^j}\cdot S) - \omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)\right| d\mu_0 \\
&+ \int_K \left|J_S\varphi_1^{v^j}(x) - J_S\varphi_1^{\bar{v}}(x)\right| \left|\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)\right| d\mu_0
\end{aligned}$$

Now, using the embedding $W \hookrightarrow C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$

$$\begin{aligned}
&\left|\left((\varphi_1^{v^j})_{\#}\mu_0 - (\varphi_1^{\bar{v}})_{\#}\mu_0|\omega\right)\right| \\
&\leq \left(\int_K |J_S\varphi_1^{v^j}(x)|d\mu_0\right) \|\omega\|_{1,\infty}\|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} + C\|(\varphi_1^{v_j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty}\|\omega\|_\infty \\
&\leq C'\|(\varphi_1^{v_j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty}\|\omega\|_W.
\end{aligned}$$

Taking supremum over all $\omega \in W$ with $\|\omega\|_W \leq 1$, we obtain that

$$\|(\varphi_1^{v^j})_{\#}\mu_0 - (\varphi_1^{\bar{v}})_{\#}\mu_0\|_{W^*} \leq C'\|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} \to 0$$

as $j \to \infty$. Combining this with lower semicontinuity of $v \mapsto \|v\|^2_{L^2([0,1],V)}$, we finally obtain that

$$E(\bar{v}) \leq \liminf_{j\to\infty} E(v^j)$$

and hence $\bar{v}$ is a global minimizer.

□

As a result of Theorem 5, one can obtain the necessary and sufficient conditions on the kernels of W and V for the two embedding assumptions of Theorem 9 to hold. In our context, in order to get $W \hookrightarrow C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ for instance, it is enough to assume that ρ and γ are C^2 functions such that all derivatives of ρ up to order 2 vanish as $x \to +\infty$. Under these assumptions, $k_W((x,T),\cdot) = \rho(|x-\cdot|^2)\gamma(\langle T, \cdot\rangle) \in C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ for all $(x,T) \in \mathbb{R}^n \times \widetilde{G}_d^n$, therefore we have $W \hookrightarrow C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ from Theorem 5.

As an important note, the formulation of (2.1) extends registration of submanifolds or rectifiable subsets in the sense that if $\mu_0 = \mu_{X_0}$ and $\mu_{tar} = \mu_{X_{tar}}$ for two oriented d-rectifiable subsets of $\mathbb{R}^n$ then (2.1) becomes equivalent, thanks to Proposition 2, to registering rectifiable subsets, i.e. to the problem (1.31).

2.1.2 General optimality conditions

Since the existence of solutions is guaranteed by Theorem 9, a natural question that we address in this section is to derive the necessary optimality conditions and the general form of those solutions. Those optimality equations will be derived in a simpler form and fashion in the discrete case and that this section can be skipped on first read. As we discussed in section 1.2.2, these are provided by the Pontryagin Maximum Principle (PMP) stated in Theorem 7. We follow the same setting as well as related works such as [79] by first rewriting the above problem as an optimal control problem on diffeomorphisms, i.e.

$$\underset{v\in L^2([0,1],V)}{\operatorname{argmin}} \left\{ \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + g(\varphi_1^v) \mid \text{s.t. } \dot{\varphi}_t^v = v_t \circ \varphi_t^v \right\}$$

with $g(\varphi_1^v) \doteq \lambda \|(\varphi_1^v)_{\#}\mu_0 - \mu_{tar}\|_{W^*}^2$. The state variables are now given by the deformations φ_t^v which we view as elements of the Banach space $\mathcal{B} \doteq \mathrm{id} + C_0^1(\mathbb{R}^n, \mathbb{R}^n)$. Let us denote, for $\phi \in \mathrm{Diff}(\mathbb{R}^n)$, $\xi_\phi : V \to C_0^1(\mathbb{R}^n, \mathbb{R}^n)$ the mapping $v \mapsto v \circ \phi$. We then introduce the Hamiltonian functional $H : C_0^1(\mathbb{R}^n, \mathbb{R}^n)^* \times \mathcal{B} \times V \to \mathbb{R}$ defined by:

$$H(p, \phi, v) = (p|v \circ \phi) - \frac{1}{2}\|v\|_V^2 \tag{2.2}$$

where p is the costate variable which is a vector distribution of $C_0^1(\mathbb{R}^n, \mathbb{R}^n)^*$ and $(p|v\circ\phi)$ denotes the duality bracket in $C_0^1(\mathbb{R}^n, \mathbb{R}^n)^*$. With the assumptions of Theorem 9, it follows from the maximum principle stated in Theorem 7 that if (v_t, φ_t^v) is a global minimum of the optimal control problem, there exists a path of costates

$p \in H^1([0,1], C_0^1(\mathbb{R}^n, \mathbb{R}^n)^*)$ such that the following equations hold:

$$\begin{cases} \dot{\varphi}_t^v = \partial_p H(p_t, \varphi_t^v, v_t) \\ \dot{p}_t = -\partial_\phi H(p_t, \varphi_t^v, v_t) \\ \partial_v H(p_t, \phi_t^v, v_t) = 0 \end{cases} \tag{2.3}$$

with the end time boundary conditions $p_1 = -\partial_\phi g(\varphi_1^v)$. From the last equation in (2.3), we can attempt to deduce the form of the optimal v. Recall the Riesz isometry operator $\mathbf{K}_V : V^* \to V$ and its inverse $\mathbf{L}_V = \mathbf{K}_V^{-1} : V \to V^*$, we get:

$$\xi^*_{\varphi_t^v} p_t - \mathbf{L}_V v_t = 0 \;\Rightarrow\; v_t = \mathbf{K}_V \xi^*_{\varphi_t^v} p_t. \tag{2.4}$$

One additional consequence of (2.3) is the following conservation of momentum. We first observe that for any $h \in C_0^1(\mathbb{R}^n, \mathbb{R}^n)$, we have

$$(\partial\phi(p|v \circ \phi)|h) = (p|d_{\phi(\cdot)} v \circ d_{(\cdot)}\phi(h)).$$

Then from (2.3) and the equation above, we can see that

$$\begin{aligned} \frac{d}{dt}(p_t|d_{(\cdot)}\varphi_t^v(u)) &= (\dot{p}_t|d_{(\cdot)}\varphi_t^v(u)) + (p_t|d_{(\cdot)}(v_t \circ \varphi_t^v)(u)) \\ &= -(\partial_\phi|_{\phi=\varphi_t^v}(p_t|v_t \circ \phi)|d_{(\cdot)}\varphi_t^v(u)) + (p_t|d_{\varphi_t^v(\cdot)} v_t \circ d_{(\cdot)}\varphi_t^v(u)) \\ &= 0. \end{aligned}$$

Therefore, we can obtain the conservation of momentum:

$$(p_t|d\varphi_t^v u) = (p_0|u), \tag{2.5}$$

for all $u \in C_0^1(\mathbb{R}^n, \mathbb{R}^n)$ and $t \in [0,1]$. Note that (2.3), (2.4) and (2.5) are generic to the LDDMM model and so far independent of the nature of the deformed objects and of the term $g(\varphi_1^v)$ in the cost. This dependency is entirely encompassed by the boundary condition $p_1 = -\partial_\phi g(\varphi_1^v)$ which we may describe a little more precisely based on the following:

Proposition 5. *The end-time momentum p_1 is a vector distribution in $C^1_0(\mathbb{R}^n,\mathbb{R}^n)^*$ of the form*

$$\begin{aligned}(p_1|u) = &\int_{\mathbb{R}^n} \alpha(x)\cdot u(x)\ d|\mu_0|(x)\\ &+\int_{\mathbb{R}^n\times\widetilde{G}^n_d} \beta(x,T)\, du|_T(x)\ d\mu_0(x,T)\\ &+\int_{\mathbb{R}^n\times\widetilde{G}^n_d} \gamma(x,T)\, div_T u(x)\ d\mu_0(x,T)\end{aligned}$$

where $\alpha:\mathbb{R}^n\to\mathbb{R}^n$, $\beta:\mathbb{R}^n\times\widetilde{G}^n_d\to(\mathbb{R}^{n\times d})^*$ and $\gamma:\mathbb{R}^n\times\widetilde{G}^n_d\to\mathbb{R}$ are continuous functions and for all $T\in\widetilde{G}^n_d$, $div_T u$ and $du|_T$ denote the divergence and differential of u restricted to T.

A condensed proof of this proposition can be found in the Appendix, although we have left aside the technical derivations related to differential calculus on the Grassmannian (this will be discussed further in Section 2.2 in the discrete setting). This result extends in a way first variation formulas for varifolds proved in [45, 55] which considered variations of rectifiable varifolds resulting from variations of the underlying rectifiable sets. This corresponds to the special case in which $\mu_0=\mu_{X_0}$. In that case, one can show, after some derivations, that the above expression of p_1 can be rewritten in the form of a vector distribution $u\mapsto\int_{\varphi^v_1(X_0)}u(x)\cdot h(x)d\mathcal{H}^d$ in $C^0_0(\mathbb{R}^n,\mathbb{R}^n)^*$ with vectors $h(x)$ normal to $\varphi^v_1(X_0)$ at each x. In our more general situation, this is however not possible and p_1 is a priori a distribution that involves first order derivatives of the test function u.

Now, the conservation law of (2.5) gives that for all $t\in[0,1]$,

$$(p_t|d\varphi^v_t u) = (p_1|d\varphi^v_1 u) = (p_0|u).$$

Using the expression of p_1 in Proposition 5, and grouping all 0-th and 1-st order terms in the resulting expressions, we may write p_t in the general form:

$$(p_t|u) = \int_{\mathbb{R}^n\times\widetilde{G}^n_d}\alpha_t(x,T)\cdot u(x)\ d\mu_0(x,T) + \int_{\mathbb{R}^n\times\widetilde{G}^n_d} B_t(x,T)d_x u|_T\ d\mu_0(x,T)$$

where $\alpha_t : \mathbb{R}^n \times \widetilde{G}_d^n \to \mathbb{R}^n$ and $B_t : \mathbb{R}^n \times \widetilde{G}_d^n \to (\mathbb{R}^{n\times d})^*$ are continuous fields, with $\alpha_1(x,T) = \alpha(x)$ and $B_1(x,T)du|_T(x) = \beta(x,T)du|_T(x) + \gamma(x,T)\mathrm{div}_T u(x)$. Furthermore, optimal vector fields satisfy $v_t = \mathbf{K}_V \xi^*_{\varphi_t^v} p_t$ and we have

$$\begin{aligned}(\xi^*_{\varphi_t^v} p_t | u) &= (p_t | u \circ \varphi_t^v) \\ &= \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha_t(x,T) \cdot u(\varphi_t^v(x))\ d\mu_0(x,T) + \int_{\mathbb{R}^n \times \widetilde{G}_d^n} B_t(x,T) d_{\varphi_t^v(x)} u|_{d_x\varphi_t^v \cdot T}\ d\mu_0(x,T).\end{aligned}$$

Denoting $K_V : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ the reproducing kernel of V, the reproducing kernel property implies that for all $u \in V$ and $x, h \in \mathbb{R}^n$, $u(x)\cdot h = \langle K_V(x,\cdot)h, u\rangle_V$. Moreover, the derivative reproducing property (1.2) gives that for any $h, h' \in \mathbb{R}^n$,

$$d_x u(h) \cdot h' = \langle \partial_1 K_V(x,\cdot)(h) \cdot h', u\rangle_V.$$

Then, we rewrite the linear maps B_t as $B_t(x,T)H = \sum_{i=1}^d b_{t,i}(x,T) \cdot H_i$ for any $H = (H_1, \ldots, H_d) \in \mathbb{R}^{n\times d}$ and where $b_i(x,T) \in \mathbb{R}^n$ are the component vector fields of B_t. By the above and the linearity of $\mathbf{K}_V$, we obtain the following general expression for optimal vector fields

$$\begin{aligned}v_t = &\int_{\mathbb{R}^n \times \widetilde{G}_d^n} K_V(\varphi_t^v(x),\cdot)\alpha_t(x,T)\ d\mu_0(x,T) \\ &+ \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \left(\sum_{i=1}^d \partial_1 K_V(\varphi_t^v(x),\cdot)(d_x\varphi_t^v(t_i)) \cdot b_{t,i}(x,T) \right) d\mu_0(x,T). \qquad (2.6)\end{aligned}$$

In contrast with LDDMM registration of submanifolds or point clouds, the expression of optimal deformation fields involves in general both the kernel function and its first order derivatives. We do not explicit the vector fields α and b_i at this point, it will be specified below in the discrete setting, see Section 2.2.2.

2.2 Numerical considerations

Having introduced a variational formulation for the varifold registration problem, we now turn more specifically to the numerical implementation of methods for solving those problems. The previous derivations were so far conducted for completely general

measures in the space $\mathcal{V}_d$. However, in practice, most geometric data sets are stored in discrete formats which are natural be represented as discret varifolds. Moreover, general varifolds can be approximated by discrete varifolds and we will study such approximation problems in Chapter 3. Therefore, in the numerical perspective, we shall focus on the the case of discrete varifolds. The first hurdle, which we start by addressing in Section 2.2.1, is to define an adequate framework for representing and computing with elements of the oriented Grassmannian.

2.2.1 Frame representation for computation

In order to come up with a computationally effective representation of discrete varifolds, we make use of the Plücker embedding introduced in section 1.3.1 to design a convenient way to represent any N Dirac masses in $\mathcal{V}_d$ as a state in $\mathbb{R}^{Nn(d+1)}$.

Let μ be a discrete varifold of the form $\mu = \sum_{i=1}^{N} r_i \delta_{(x_i, T_i)}$. The main idea is to represent, for each i, the weight r_i and oriented tangent space T_i together by a frame which spans T_i. To be precise, for each i, we can find a frame $\{u_i^{(1)}, \cdots, u_i^{(d)}\}$ such that

$$T_i = \frac{u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}}{|u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}|} \text{ and } r_i = |u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}|. \tag{2.7}$$

In other words, the oriented space spanned by the frame $\{u_i^{(1)}, \cdots, u_i^{(d)}\}$ corresponds to T_i while its d-volume matches the weight r_i. Given such a choice of frame for each i, we can then identify μ with the state variable $q = (x_i, u_i^{(1)}, \cdots, u_i^{(d)})_{i=1,\cdots,N}$ in the vector space $\mathbb{R}^{Nn(d+1)}$. Of course such representation is not unique since for each i, we can always find another frame satisfying (2.7). Such non-uniqueness results in additional invariances which we will study more thoroughly in the next section. Conversely, such a frame q with $(u_i^{(1)}, \cdots, u_i^{(d)})$ a matrix of rank d for all i, corresponds to the unique discrete oriented varifold defined by the relations of (2.7); we will denote it by μ^q in what follows.

In this representation, the kernel metrics for discrete varifolds expressed in (1.30)

can be explicitly written as

$$\langle \mu, \mu' \rangle^2_{W^*} = \sum_{i=1}^{N} \sum_{j=1}^{M} r_i r'_j \rho(|x_i - x'_j|^2) \gamma \left(\frac{1}{r_i r'_j} \det(u_i^{(k)} \cdot u'^{(l)}_j)_{k,l} \right) \tag{2.8}$$

where $r_i = |u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}| = \sqrt{\det(u_i^{(k)} \cdot u_i^{(l)})_{k,l}}$. Note that this expression does not depend on the choice of frames that satisfy the conditions of (2.7) for μ (and similarly for μ'). In the case where μ' is a more general non-discrete varifold in $\mathcal{V}_d$, the computation of $\langle \mu, \mu' \rangle^2_{W^*}$ involves integrals over $\mathbb{R}^n \times \widetilde{G}_d(\mathbb{R}^n)$ of the kernel functions, which requires introducing specific quadrature schemes for approximating them. We do not address those issues in more details in this work as it needs particular discussion depending on the nature, regularity and dimension of the varifolds under consideration. Provided such adequate quadrature schemes have been defined, the W^* metric then formally reduces to an expression equivalent to (2.8) in which the x'_j, u'_j and r'_j are now the quadrature nodes and associated weights of the scheme.

In practice, computations of varifold kernel metrics for different classes of kernels and gradients of the metrics can be conveniently implemented with automatic differentiation pipelines. In our MATLAB implementation, we make use of the recent KeOps library [80] which allows to generate CUDA functions for the low-level kernel sum evaluations and their automatic differentiation.

2.2.2 Discrete registration model

This frame representation also provides a convenient setting to express the diffeomorphism action and registration problem on discrete varifolds. Indeed, let φ be a diffeomorphism of $\mathbb{R}^n$ and $\mu \in \mathcal{V}_d^N$, the pushforward action $\varphi_\# \mu$ in (1.19) is equivalent to the following action in the frame model:

$$\varphi_\# q := (\varphi(x_i), d_x\varphi(u_i^{(1)}), \cdots, d_x\varphi(u_i^{(d)}))_{i=1,\cdots,N}.$$

Now, this allows us to rewrite the former infinite-dimensional optimal control problem by considering instead the finite-dimensional state variable $q \in \mathbb{R}^{Nn(d+1)}$. In the next

paragraphs, we give a direct derivation of the optimality conditions in this discrete setting, in order to arrive at simpler and more explicit equations than the general abstract derivations presented in Section 2.1.2.

Following once again the Pontryagin maximum principle approach, the Hamiltonian for this discrete representation is given by:

$$H(q,p,v) \doteq \sum_{i=1}^{N} \left[p_i^x \cdot v(x_i) + \sum_{k=1}^{d} p_i^{u_k} \cdot d_{x_i} v(u_i^{(k)}) \right] - \frac{1}{2}\|v\|_V^2 \tag{2.9}$$

with $p^x, p^{u_k} \in \mathbb{R}^n$ denoting respectively the costates for the position x and frame vector $u^{(k)}$ variables. The PMP then shows that optimal trajectories of the registration problem are governed by the dynamical system:

$$\begin{cases} \dot{x}_i(t) = v_t(x_i(t)) \\ \dot{u}_i^{(k)} = d_{x_i} v(u_i^{(k)}) \\ \dot{p}_i^x = -d_{x_i} v^T p_i^x - \sum_{k=1}^{d} d_{x_i}^{(2)} v(\cdot, u_i^{(k)})^T p_i^{u_k} \\ \dot{p}_i^{u_k} = -d_{x_i} v^T p_i^{u_k} \end{cases} \tag{2.10}$$

while optimal vector fields v satisfy

$$v_t(\cdot) = \sum_{i=1}^{N} \left(K(x_i(t),\cdot) p_i^x(t) + \sum_{k=1}^{d} \partial_1 K(x_i(t),\cdot)(u_i^{(k)}(t)) \cdot p_i^{u_k} \right). \tag{2.11}$$

Plugging the above expression of v with respect to (p,q) in the Hamiltonian (2.9) gives the reduced Hamiltonian $H_r(p,q) \doteq H(p,q,v)$ which writes:

$$\begin{aligned} H_r(p,q) = \frac{1}{2} \sum_{i,j=1}^{N} \Big[& p_i^x \cdot K(x_i,x_j) p_j^x + p_i^x \cdot \sum_{k=1}^{d} \partial_1 K(x_j,x_i)(u_j^{(k)}) \cdot p_j^{u_k} \\ & + \sum_{k=1}^{d} p_i^{u_k} \cdot \partial_2 K(x_j,x_i)(u_j^{(k)}) p_j^x + \sum_{k,l=1}^{d} p_i^{u_k} \cdot \partial_{1,2}^2 K(x_j,x_i)(u_j^{(l)}, u_i^{(k)}) p_j^{u_l} \Big] \end{aligned} \tag{2.12}$$

and (2.10) becomes a coupled system in the variables q and p which is the system of reduced Hamiltonian equations. Consequently, the set of optimal paths is entirely determined by the initial values $(q(0), p(0))$ and the value of the reduced Hamiltonian $H_r(p(t), q(t)) = \frac{1}{2}\|v_t\|_V^2$ is conserved along an optimal trajectory.

There are in addition several other conserved quantities in such a system as evidenced by the following lemma:

Lemma 1. *For any $i = 1, \dots, N$, the matrix*

$$D^i(t) \doteq \left(\langle u_i^{(k)}(t), p_i^{u_\ell}(t) \rangle \right)_{1 \le k, \ell \le d},$$

is constant in time.

Proof. Using the Hamiltonian equations written above, we have for all $k, l = 1, \dots, d$

$$\frac{d}{dt} \left(D^i(t) \right)_{k,\ell} = \langle d_{x_i} v(u_i^{(k)}(t)), p_i^{u_\ell} \rangle - \langle u_i^{(k)}, d_{x_i} v^T p_i^{u_\ell} \rangle = 0.$$

Hence $D^i(t)$ is a constant matrix. □ □

Note that, at this point, all those equations are fundamentally modelling the deformation of the frames $\{x_i, (u_i^{(k)})\}$ but are not yet taking into account the invariances that result from the representation of the discrete oriented varifolds as oriented frames. Those extra invariances can be derived from the boundary conditions of the PMP:

$$p(1) = -\partial_q g(q)|_{q=q(1)}, \text{ with } g(q) = \lambda \|\mu^q - \mu_{tar}\|^2_{W^*}. \tag{2.13}$$

As a clear consequence of (2.7), μ^q and thus $g(q)$ are independent of the choices of the frame vectors $(u_i^{(k)})_{k=1,\dots,d}$ that span the same oriented vector spaces T_i with the same d-volumes r_i. This in turn leads to a set of conditions satisfied by the different components of the final costate $p(1)$ and, with Lemma 1, of the full path $p(t)$. These are summed up by the following result:

Proposition 6. *Let $(q(t), p(t))$ be optimal trajectory, then for all i, the matrices $D^i(t)$ as defined above are constant scalar matrices. In particular, we have $p_i^{u_k}(t) \perp Span(\{u_i^{(\ell)}(t)\}_{\ell \neq k})$ for all $t \in [0,1]$, $i = 1, \dots, N$ and $k = 1, \dots, d$.*

Proof. We can treat the case of each particle i separately and thus, without loss of generality, we may directly assume that $N = 1$. We write $q(t) = (x(t), u^{(1)}(t), \cdots, u^{(d)}(t))$,

$p(t) = (p^x(t), p^{u_1}(t), \ldots, p^{u_d}(t))$ for the state and costate variables along an optimal trajectory and

$$U \doteq \text{Span}\{u^{(1)}(1), \cdots, u^{(d)}(1)\}.$$

Consider the group of linear transformations, $G \doteq \text{SL}(U) \oplus \text{GL}(U^\perp)$, i.e., for any $\text{g} \in G$,

$$\text{g}(x) = \text{g}_{/\!/}(x_U) + \text{g}_\perp(x_{U^\perp}),$$

where x_U and $x_{U^\perp}$ are the orthogonal projections of x on U and $U^\perp$, with $\text{g}_{/\!/} \in \text{SL}(U)$ and $\text{g}_\perp \in \text{GL}(U^\perp)$. The Lie algebra of G is $\mathfrak{g} = \mathfrak{sl}(U) \times \mathcal{L}(U^\perp)$ and $\mathfrak{sl}(U)$ is the set of all zero trace linear transformations of U. Now, consider the action of G on $\mathbb{R}^{(d+1)n}$ defined as:

$$\text{g} \cdot q := (q_0, \text{g}(q_1), \cdots, \text{g}(q_d)).$$

for any $q = (q_0, \ldots, q_d) \in \mathbb{R}^{(d+1)n}$. We see that $\mu^{\text{g}\cdot q(1)} = \mu^{q(1)}$ for all $\text{g} \in G$ and therefore $g(\text{g} \cdot q(1)) = g(q(1))$.

Now, if we let $\{\text{g}_t\}$ be a smooth curve in G that satisfies $\text{g}_0 = id$ and $\frac{d}{d\tau}|_{\tau=0}\text{g}_\tau = h \in \mathfrak{g}$, differentiating the equality $g(\text{g}_\tau \cdot q(1)) = g(q(1))$ shows that for any $h \in \mathfrak{g}$, we have

$$0 = (p(1)|h \cdot q(1)) = \sum_{k=1}^{d} \langle p^{u_k}(1), h(u^{(k)}(1))\rangle$$

Since $h \in \mathfrak{g}$, we must have that $h|_U$ is a zero trace linear map. For any $1 \leq i < j \leq d$, we may choose h such that $h(u^{(i)}(1)) = -h(u^{(j)}(1))$ and $h(u^{(k)}(1)) = 0,\ \forall k \notin \{i, j\}$, which leads to $\langle u^{(i)}(1), p^{u_i}(1)\rangle = \langle u^{(j)}(1), p^{u_j}(1)\rangle$. Consequently,

$$\langle u^{(1)}(1), p^{u_1}(1)\rangle = \cdots = \langle u^{(d)}(1), p^{u_d}(1)\rangle = \alpha$$

for some constant α. In addition, for any $i \neq j$, we can also choose h such that $h(u^{(i)}(1)) = u^{(j)}(1)$ and $h(u^{(k)}(1)) = 0,\ \forall k \notin \{i, j\}$, which gives $\langle u^{(i)}(1), p^{u_j}(1)\rangle = 0$. It results that $D(1) = \alpha.I_{d\times d}$.

Finally, since $D(t)$ is constant by Lemma 1, we obtain that

$$D(t) = \begin{pmatrix} \alpha & & 0 \\ & \ddots & \\ 0 & & \alpha \end{pmatrix},$$

for all $t \in [0, 1]$. □

This result is particularly interesting from a computational point of view as it allows to partly alleviate the redundancy introduced by the frame representation of Grassmannians. Indeed, we see that the costates $p(t)$ actually lie in affine subspaces of $\mathbb{R}^{Nn(d+1)}$ of lower dimensions $N(n + d(n - d) + 1)$, which is precisely the dimension of the 'true' state space $(\mathbb{R}^n \times \widetilde{G}_d(\mathbb{R}^n) \times \mathbb{R})^N$.

2.2.3 Registration algorithm

With the optimality equations of the previous section, we can now easily design an algorithm based on the *geodesic shooting* method discussed in section 1.2.2 to solve the discrete registration problem. Let us now explain in detail about this algorithm for our particular problem. As mentioned earlier, optimal trajectories are completely determined, through the Hamiltonian equations (2.10) and (2.11), by the initial conditions $q(0) = q_0$, which is known, and $p(0)$. We can then optimize the cost function over $p(0)$.

The main issue is to compute the gradient of the total energy E with respect to the initial costate $p(0)$. The regularization term being equal to $H_r(p(0), q(0))$ thanks to the conservation of the reduced Hamiltonian, its gradient can be obtained by direct differentiation of (2.12). The fidelity term $g(q(1))$ on the other hand depends indirectly on the initial costate $p(0)$ via the integration of the forward reduced Hamiltonian equations. The gradient of $g(q(1))$ with respect to $p(0)$ can be computed by flowing backward in time the adjoint Hamiltonian system

$$\dot{Z}(t) = -dF(q(t), p(t))^T Z(t) \tag{2.14}$$

where $F(q,p) = (\partial_p H_r(p,q), -\partial_q H_r(p,q))$, $Z(t) = (\tilde{q}(t), \tilde{p}(t)) \in \mathbb{R}^n \times (\mathbb{R}^n)^d$ the adjoint variables of the system, together with the end-time conditions $\tilde{q}(1) = -\partial_q g(q)|_{q=q(1)}$ and $\tilde{p}(1) = 0$. Although being a linear system of ODEs, the adjoint equations can be tedious to derive and implement, in particular given the rather intricate expression of the reduced Hamiltonian function considered here. Instead, the differential appearing on the right hand side of (2.14) can be approximated efficiently based on the finite difference trick proposed in [30] (Section 4.1). Indeed, it can be rewritten as follows:

$$dF(q,p)^T Z = \begin{pmatrix} \partial_p(\partial_q H_r) \cdot \tilde{q} - \partial_q(\partial_q H_r) \cdot \tilde{p} \\ \partial_p(\partial_p H_r) \cdot \tilde{q} - \partial_q(\partial_p H_r) \cdot \tilde{p} \end{pmatrix}$$

which only involves directional derivatives of the components of the function F in the directions of $\tilde{q}$ and $\tilde{p}$. We then specifically approximate the above by centered finite difference

$$dF(q,p)^T Z \approx \begin{pmatrix} \alpha \\ -\beta \end{pmatrix}, \quad \text{with} \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{F(q - \epsilon\tilde{p}, p + \epsilon\tilde{q}) - F(q + \epsilon\tilde{p}, p - \epsilon\tilde{q})}{2\epsilon}$$

for some small $\epsilon > 0$, which only requires at each time t two evaluations of the same function F that appears in the forward reduced Hamiltonian equations.

With the above approach to compute the gradient with respect to $p(0)$, the registration algorithm then consists of essentially the same steps as the standard LDDMM algorithm for submanifolds:

1: **repeat**
2: From $(q(0), p(0))$ compute $(q(t), p(t))$ by forward integration of the reduced Hamiltonian system given by (2.10) and (2.11).
3: Compute $g(q(1))$ and $-\partial_q g(q)|_{q=q(1)}$.
4: Integrate backward the adjoint Hamiltonian equations (2.14) to obtain $\partial_{p(0)} g(q(1))$.
5: Deduce the gradient of the full cost function with respect to $p(0)$.
6: Update $p(0)$.
7: **until** convergence

For the numerical ODE integration steps of lines 2 and 4, we use a standard RK4 scheme with regular time samples in $[0, 1]$, where we typically take $T = 15$

time steps in most of the experiments that we present in the next section. Note that one can easily replace the RK4 scheme by even higher order or adaptive step methods although in practice we have found this to be unnecessary for the types of ODEs involved here. The optimization update in line 6 follows the limited memory BFGS algorithm, specifically the implementation provided by the HANSO library [81]. One can further take additional advantage of the dimensionality reduction provided by Proposition 6 by restricting each of the components $p_i^{u_k}(0)$ to the linear subspace $\mathrm{Span}(\{u_i^{(\ell)}(0)\}_{\ell \neq k})^{\perp}$. Lastly, as in Section 2.2.1, all kernel summation and differentiation operations appearing in both the varifold fidelity terms and Hamiltonian equations are coded in CUDA using the KeOps library [80]. The full implementation of the varifold approximation and diffeomorphic registration approach is available at https://github.com/charoncode/Var_LDDMM together with the scripts and data of some of the simulations presented in the next section.

2.3 Numerical results

We now present some results of the previous algorithms on discrete varifolds of dimension $d = 1$ and $d = 2$. In all these experiments, we choose the deformation kernel K of V to be a diagonal Gaussian kernel $K(x, y) = \exp(-\frac{|x-y|^2}{\sigma_V^2})\mathrm{id}$. The kernel function ρ is a Gaussian of scale σ_ρ. The choice of these scales is adapted to the sizes of the shapes in each of the experiment. The function γ is chosen, depending on the situation, in the different classes of functions discussed in detail in [55], the main distinction being whether the considered varifolds are rectifiable or not according to the conditions given by Theorem 10 and Theorem 11 and whether the shapes carry a relevant orientation or not. In particular, one can use $\gamma(t) = t^2$ to recover an orientation invariant fidelity metric, or $\gamma(t) = e^{-\frac{2}{\sigma_\gamma^2}(1-t)}$ which leads to an orientation sensitive distance that satisfy the conditions of Theorem 11. All simulations are run on a desktop computer equipped with a NVIDIA Quadro P5000 graphics card.

2.3.1 1D example

We begin with a toy example of standard curve matching to compare the result and performance of our discrete varifold LDDMM registration algorithm with the state-of-the-art LDDMM approach for curves such as the implementations of [45, 55]. Recall that in the previous works, the matching problem was formulated as (1.31) and varifold metrics were used as fidelity terms. The essential difference being that the state of the optimal control problem is there the set of vertices of the deformed template curve which is only converted to a varifold for the evaluation of the fidelity term at each iteration. But the dynamics of geodesics still correspond to usual point set deformation under the LDDMM model.

For the varifold LDDMM approach, we consider the pushforward action of diffeomorphism (1.19) which we have seen in Proposition 2 to be compatible with the action of diffeomorphisms on curves. Therefore, the two formulations and optimization problems are theoretically equivalent up to the discretization precision. We verify it with the example of Figure 2-1 for which both algorithms are applied with the same deformation kernel, varifold metric and optimization scheme. Note that in the varifold registration model, template and target curves are first (and only once at the beginning) converted to their discrete varifold representations as explained in the end of Section 1.3.1.

As we can see, the resulting geodesics and deformations are consistent between the two methods. This is also corroborated by the very similar values of the energy at convergence. Interestingly however, although each iteration in our model is arguably more expansive numerically compared to standard curve-LDDMM due to the increased complexity of the Hamiltonian equations, the algorithm converges in a significantly lesser number of iterations.

Figure 2-2 shows an example of matching on more general discrete varifolds that

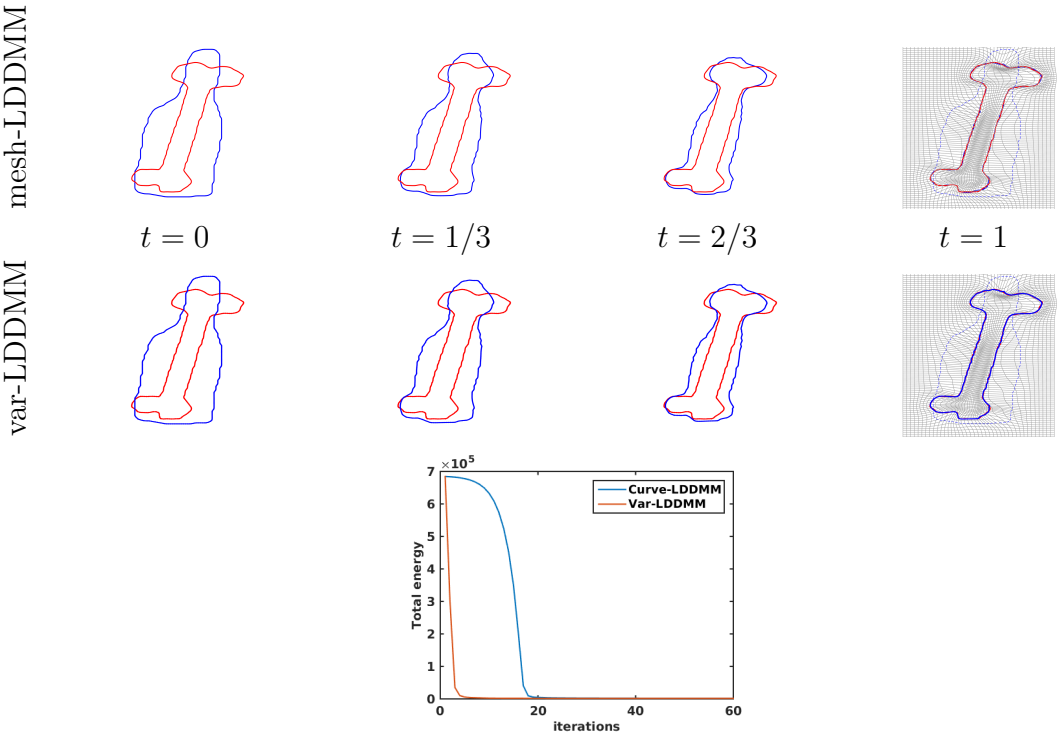


Figure 2-1. Curve registration using point-mesh LDDMM (1st row) and our proposed discrete varifold LDDMM (2nd row). On the last row is shown the evolution of the total energy across the iterations for both algorithms.

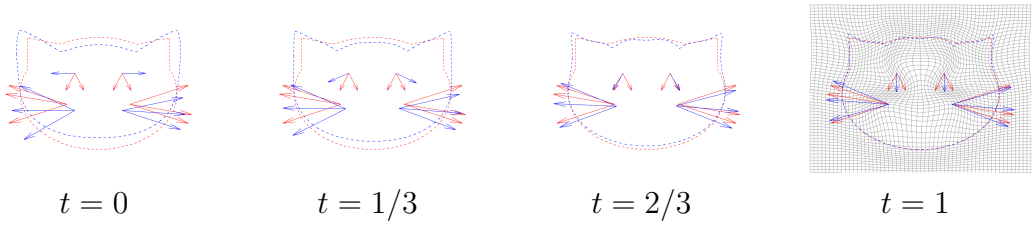


Figure 2-2. Registration of multi-directional sets. The lengths of vectors correspond to the weights of the Dirac varifolds.

involve varying number of directions at different spatial locations. This example used an oriented Gaussian kernel for the fidelity term. Although purely synthetic, it illustrates the potentialities of the proposed approach to register data with complex directional patterns.

2.3.2 2D example

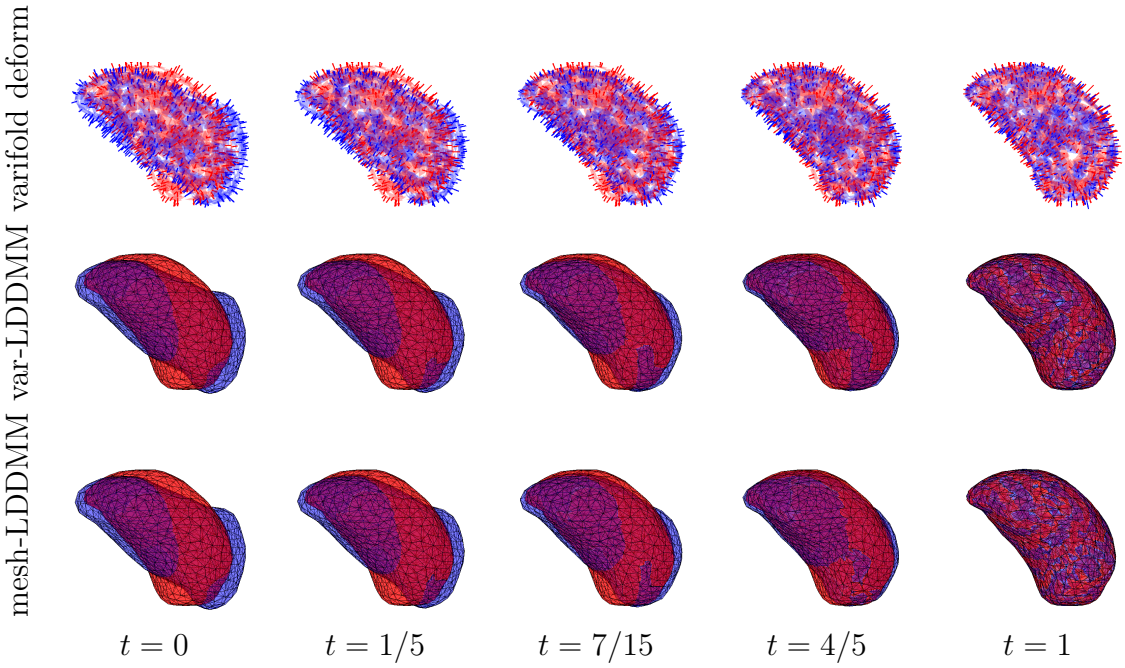


Figure 2-3. Surface registration of two amygdalas (data courtesy of S. Ardekani) using discrete varifold LDDMM (1st and 2nd row) and surface mesh LDDMM (3rd row). The first row depicts the evolution of the deformed tangent spaces along the geodesic. The parameters used are the same for both methods; namely a weighting constant $\lambda = 10$ between the regularization and fidelity term, a deformation scale $\sigma_V = 4.75$, a scale $\sigma_\rho = 3$ for the spatial kernel of the fidelity term and a Gaussian function on the sphere of scale $\sigma_\gamma = 1$ for the function γ.

First, as a sanity check, we compare our 2-varifold registration approach applied to triangulated surfaces with the previous LDDMM mesh surface matching implementation of [45, 55] using the same kernel size parameters, in which case we expect both approaches to be theoretically equivalent as pointed out in the last paragraph of Section 2.1.1. Shown in Fig. 2-3 are triangulated surfaces of amygdala segmented from

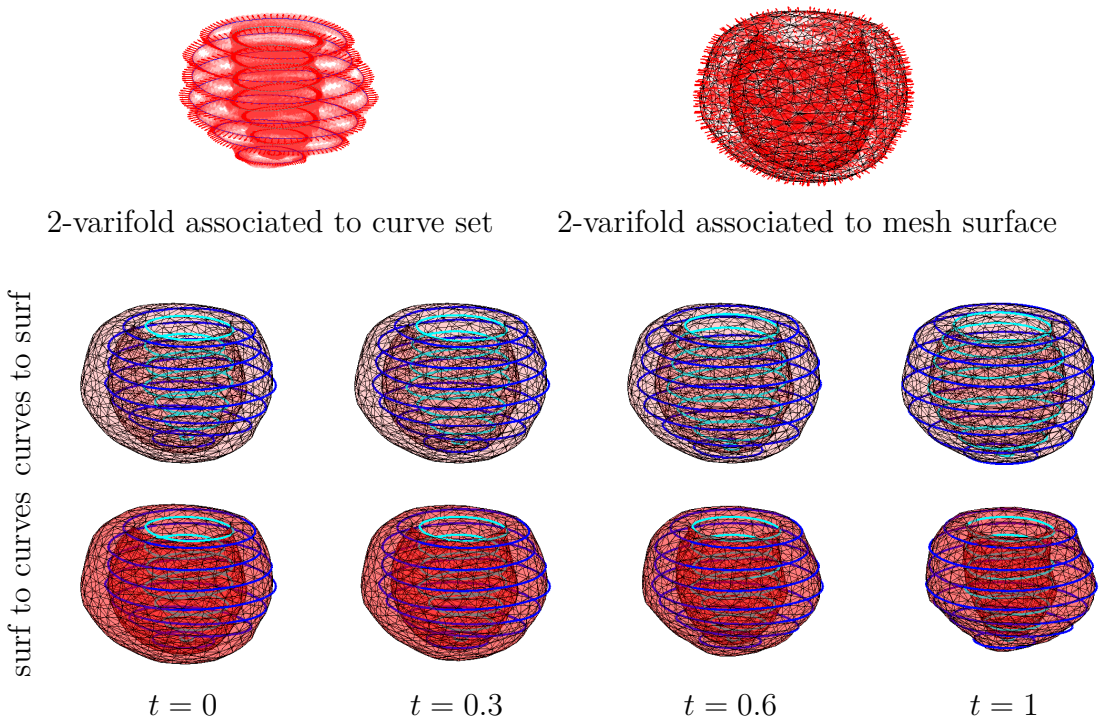


Figure 2-4. Registration between two shapes of hearts of different nature. On top row: illustration of the 2-varifolds associated to the sectional contour curves (left) for a first subject and to a triangulated surface (right) for the second subject. In the second and third rows are shown the two results of varifold registration of surface to contour curves and contour curves to surface respectively.

two different subjects of the BIOCARD database [82], containing 563 vertices, 1122 triangles and 488 vertices, 972 triangles respectively. Following the simple procedure outlined at the beginning of Section 3.2.1, we obtain discrete 2-varifolds (one Dirac for each triangle). The first row in the figure shows the optimal deformation estimated with our approach through the evolution of the discrete varifold of the source shape (red) to the target varifold (blue). Discrete varifolds are here displayed in the form of tangent patches and normal vectors (instead of 2-frames) for the purpose of better visualization. Now, the estimated vector fields v_t define a path of dense deformations of the full space which we can also apply to deform the original triangulated surface, which we show on the second row of Fig. 2-3. This is very comparable to the result of the surface mesh LDDMM registration approach displayed on the third row. In terms

of computation times, the varifold registration takes a total of 494s (0.99s per iteration of BFGS) against 92.5s (0.18s per iterations) for the surface LDDMM algorithm. This difference comes from mainly two factors: the fact that the numerical complexities are quadratic in the number of Diracs (i.e. triangles) for varifold matching as opposed to the number of vertices for surface LDDMM, and from the increased dimensionality of the Hamiltonian systems in our model.

In Fig. 2-4, we consider a more challenging registration scenario which was originally studied in [83]. Here, one of the two shape is a triangulated surface of a heart membrane segmented from high resolution CT imaging while the second one only consists of a sparse set of cross-sectional curves of the heart contour obtained from lower resolution clinical cardiac MRI data. The varifold framework of this paper leads to an alternative registration approach to the one proposed in [83] that relies on a tailored closest point fidelity cost for the surface to curve set comparison. In our case, we instead represent both shapes as 2-varifolds and register them using the exact same varifold registration algorithm as in the previous example. The triangulated surface is again associated to a discrete 2-varifold in the same way as above. As for the set of cross-sectional curve set, we first obtain its 1-varifold representation $\{x_i, u_i^{(1)}\}$ which involve the tangent vectors $u_i^{(1)}$ to the curve that passes through x_i. We then complete it into a 2-varifold by adding a second "vertical" (i.e. inter-sectional) frame vector $u_i^{(2)}$, which can be estimated in this case by simply finding the projection of x_i onto the corresponding curve in the section immediately above (note that this does involve any attempt to estimate an actual surface mesh of the data). We show the 2-varifolds associated to each shape in the first row of Fig. 2-4 and as well as the result of the 2-varifold registration both from curve set to surface and surface to curve set. In each case, we have again applied the estimated deformation between varifolds on the original shapes for visualization.

Along the same lines, we finally look into the case of even less structured data

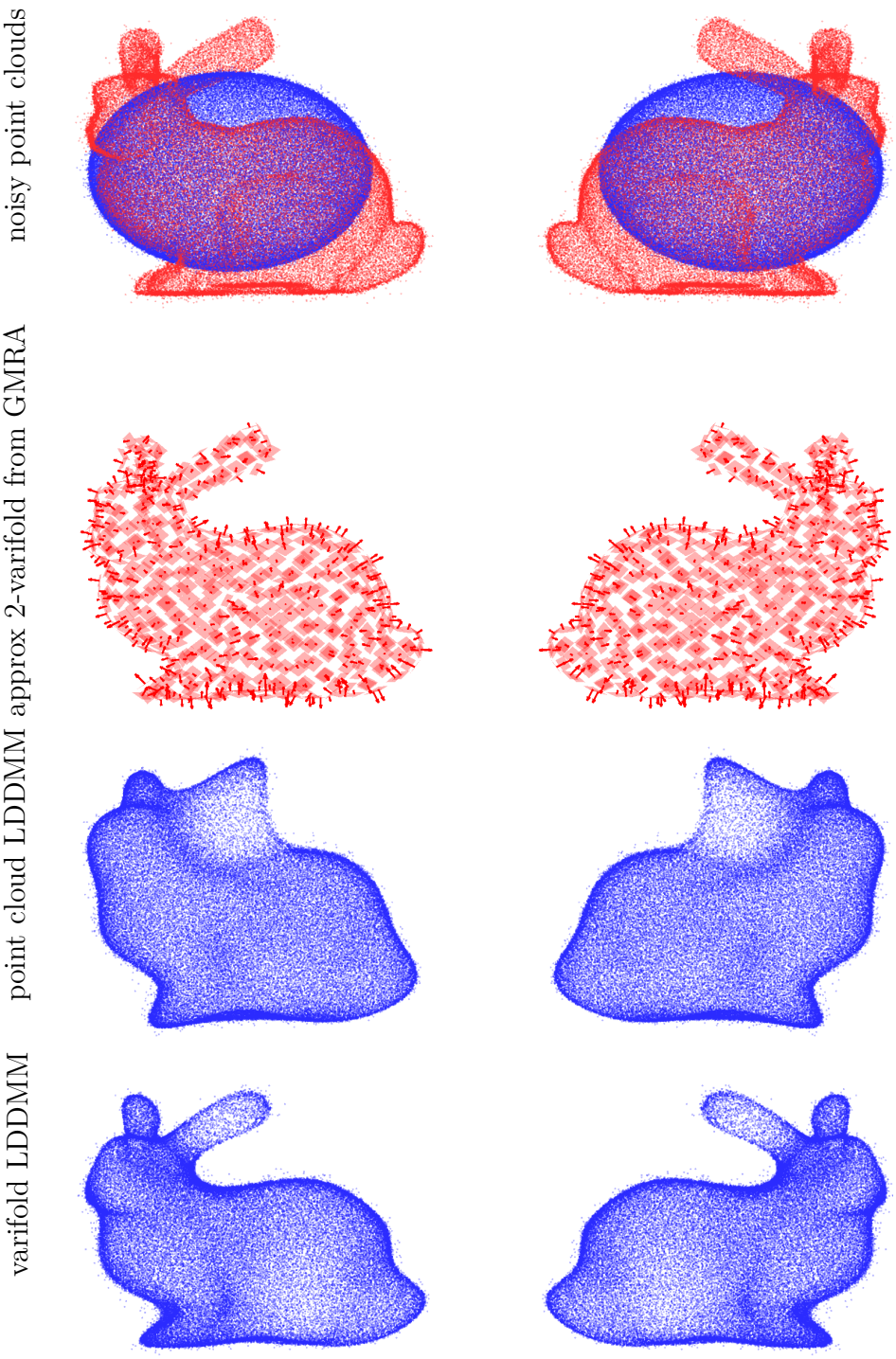


Figure 2-5. Registration of noisy point clouds. Top row: the source (blue) and target (red) point clouds with respectively 58962 and 54834 points. Second row: illustration of the target 2-varifold obtained by GMRA with only 512 Diracs. Third row: result of direct registration of the raw point clouds. Bottom row: registration estimated from the approximate 2-varifolds.

objects. Specifically, as displayed on the first row of Fig. 2-5, we consider two noisy point clouds which are obtained by first randomly selecting vertices from the groundtruth surfaces (with replacement) and then adding some Gaussian noises ($\sigma = 0.028$) to the position of each sampled point. A first possible registration approach could be to treat such point clouds as standard measures of $\mathbb{R}^3$ (i.e. 0-varifolds) and follow the simple point distribution LDDMM algorithm for unlabelled point sets proposed in [47]. The result shown on the third row of Fig. 2-5 illustrates the shortcomings of such a model for this type of data. Indeed, one can see that, in the absence of any tangential information, many details of the target shape are not well-recovered. Furthermore, this point set model is not robust to sampling changes and imbalances which results in the mismatches observed below the ear region. An arguably more adequate method would be to exploit the fact that these point clouds are close to their underlying surfaces. However, due to noise and the presence of outliers, estimating triangulations of the point clouds with standard meshing algorithms can prove particularly challenging and inefficient. Instead, our approach consists in directly learning the 2-varifold structure from the point clouds based on the geometric multi-resolution analysis (GMRA) framework developed in [84]. Here, we fix a specific scale and GMRA then provides local partitions with estimates of tangent planes to the point clouds which eventually gives us an approximate representation as a 2-varifold illustrated on the second row of Fig. 2-5. Besides its robustness and numerical efficiency, such manifold learning algorithm is also particularly well suited for our proposed registration framework since it naturally leads to approximations in the form of 2-varifolds (and generally not meshes). In the last row of Fig. 2-5, we show the deformed point cloud resulting from the deformation estimated by the 2-varifold registration algorithm. It obviously outperforms the direct point cloud registration described above both in terms of quality of matching but also computation time (10 mins vs 39 mins in total).

Chapter 3

Characterization of kernel metrics and discrete approximation of varifolds

In the previous chapters, we have seen that the kernel metrics on $\mathcal{V}_d$ play an important role in registration due to their tractability. However, some important questions regarding the kernel metrics on $\mathcal{V}_d$ have not been addressed yet. In this chapter, we start by investigating sufficient conditions on such kernels to recover true metrics on the whole varifold space and examining the relationship between those kernel metrics and the classical metrics on $\mathcal{V}_d$ introduced earlier in Section 1.3.3. We then study the quantization in the space of varifolds and propose a projection-based approach for varifold compression using kernel metrics. At last, we investigate the consistency of the discretized registration problem. Results of this chapter were partially published in [77].

3.1 Characterization of kernel distances

Recall that, as discussed in section 1.3.3, we assume the kernels on $\mathbb{R}^n \times \widetilde{G}_d^n$ are RZ kernels (1.28) and d_{W^*} is a pseudo-distance between varifolds defined in (1.27). It is a natural question to ask under which conditions it leads to an actual distance. Most past works have addressed this question focusing on the case of varifolds representing

submanifolds or reunion of submanifolds [45, 55]. We can first provide an extension of these results to the general case of oriented rectifiable varifolds. A key notion for the rest of this section is the one of C_0-universality of kernels:

Definition 9. *A positive definite kernel k on a metric space $\mathcal{M}$ is called C_0-universal when its RKHS is dense in $C_0(\mathcal{M})$ for the uniform convergence topology.*

C_0-universality has been studied in great length in such works as [49, 85]. In particular, one can provide characterizations of C_0-universality for certain classes of kernels and spaces $\mathcal{M}$. In the case of translation-invariant kernels on $\mathcal{M} = \mathbb{R}^n$ for instance, it has been established that C_0-universal kernels are the ones which can be expressed through the Fourier transform of finite Borel measures with full support on $\mathbb{R}^n$, which includes: compactly-supported kernels, Gaussian kernels, Laplacian kernels... With the Definition 6, we have the following sufficient condition:

Theorem 10. *Suppose k^{pos} is a C_0-universal kernel on $\mathbb{R}^n$, $\gamma(1) > 0$ and $\gamma(t) \neq \gamma(-t)$, $\forall t \in [-1, 1]$. Let $(X, T(\cdot))$ and $(Y, S(\cdot))$ be two oriented $\mathcal{H}^d$-rectifiable sets with $\mathcal{H}^d(X)$, $\mathcal{H}^d(Y) < \infty$. If $\|\mu_X - \mu_Y\|_{W'} = 0$, then $\mathcal{H}^d(X \triangle Y) = 0$ and $T = S$ $\mathcal{H}^d$-a.e.*

Proof. We first prove that $\mathcal{H}^d(X \triangle Y) = 0$. Let us denote by W^{pos} and W^G the RKHS associated to kernels k^{pos} and k^G respectively. Suppose that X and Y are rectifiable sets as above such that $\|\mu_X - \mu_Y\|_{W^*} = 0$ and $\mathcal{H}^d(X \triangle Y) > 0$. Without loss of generality, we may assume that $\mathcal{H}^d(X \setminus Y) > 0$. From Lusin's theorem, there exists a subset U of X such that $T|_U$ is continuous and $\mathcal{H}^d(X \setminus U) < \mathcal{H}^d(X \setminus Y)$. Let us denote by $E := U \cap (X \setminus Y)$, we see that $\mathcal{H}^d(E) > 0$. Since for $\mathcal{H}^d$ *a.e.* $x \in E$,

$$\limsup_{r \to 0} \frac{\mathcal{H}^d(B_r(x) \cap E)}{\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)} r^d} \geq \frac{1}{2^d},$$

(cf [51]), there exists $x_0 \in E$, $\mathcal{H}^d(B_r(x_0) \cap E) > 0$ for any $r > 0$.

Let $g : \widetilde{G}_d^n \to \mathbb{R}$ be defined by $g(\cdot) = k^G(T(x_0), \cdot)$. Since $x \longmapsto g(T(x))$ is continuous on E and $g(T(x_0)) > 0$, there exists $r_0 > 0$ such that $\forall\ x \in B_{r_0}(x_0) \cap E,\ g(T(x)) > 0$. Let $A \doteq B_{r_0}(x_0) \cap E$ and $h(x) := \mathbf{1}_A(x)$, then $\mathcal{H}^d(A) > 0$ and $g(T(x)) > 0,\ \forall\ x \in A$. Using the density of $C_c(\mathbb{R}^n)$ in $L^1(\mathbb{R}^n, \mathcal{H}^d \llcorner (X \cup Y))$ together with the fact that k^{pos} is C_0-universal, there exist $\{f_j\}_{j=1}^\infty \subset C_c(\mathbb{R}^n)$ and $\{h_j\}_{j=1}^\infty \subset W^{pos}$ such that $\lim_{j\to\infty} f_j = h$ in $L^1(\mathbb{R}^n, \mathcal{H}^d \llcorner (X \cup Y))$ and $\|f_j - h_j\|_\infty < \frac{1}{j}$. Now, since $h_j \otimes g \in W$ and $\mu_X = \mu_Y$ in W^*, we have

$$
\begin{aligned}
0 &= (\mu_X - \mu_Y)(h_j \otimes g) \\
&= \int_X h_j(x) g(T(x)) d\mathcal{H}^d(x) - \int_Y h_j(y) g(S(y)) d\mathcal{H}^d(y) \to \int_A g(T(x)) d\mathcal{H}^d(x) > 0,
\end{aligned}
$$

which is a contradiction. Hence we have $\mathcal{H}^d(X \triangle Y) = 0$

Next, we show that $T(x) = S(x)$ $\mathcal{H}^d$-*a.e.*. Let $F := \{x \in X | T(x) = -S(x)\}$ and assume that $\mathcal{H}^d(F) > 0$. From Lusin's theorem, there exists subset $F' \subset F$ such that $T|_{F'}$ is continuous and $\mathcal{H}^d(F') > 0$. Using the upper density argument as above, we can find $z_0 \in F'$ such that $\mathcal{H}^d(B_r(z_0) \cap F') > 0$ for all $r > 0$. Since the map $x \mapsto \langle T(x), T(z_0) \rangle$ restricted to F' is continuous, there exists a $\delta_0 > 0$ satisfying:

$$
\langle T(x), T(z_0) \rangle > 0,\ \forall x \in B_{\delta_0}(z_0) \cap F'.
$$

Define $B := B_{\delta_0}(z_0) \cap F'$, $\eta(\cdot) := \gamma(\langle \cdot, T(z_0) \rangle)$ and $u(x) := \eta(T(x)) - \eta(S(x))$. Observe that, from the assumption $\gamma(t) \neq \gamma(-t),\ \forall t \in [-1, 1]$,

$$
u(x) = \eta(T(x)) - \eta(-T(x)) \neq 0,\ \forall x \in F'.
$$

From this, we may assume that $u(x) > 0,\ \forall x \in F'$. Let $\{f_j'\}_j$ and $\{h_j'\}_j$ be sequences in $C_c(\mathbb{R}^n)$ and W_{pos} such that f_j' converges to $\mathbf{1}_B$ in $L^1(\mathbb{R}^n, \mathcal{H}^d \llcorner F)$ and $\|f_j' - h_j'\|_\infty < 1/j$. We obtain

$$
0 = (\mu_X - \mu_Y | h_j' \otimes \eta) = \int_X h_j'(x) u(x) d\mathcal{H}^d(x) \to \int_B u(x) d\mathcal{H}^d(x) > 0,
$$

which is impossible. □

Note that the first part of the proof directly gives an equivalent statement for unoriented rectifiable varifolds, generalizing the result of [45].

However, the previous proposition does not necessarily lead to a distance on the whole space $\mathcal{V}_d$. For example, we can consider the case that $n = 2$ and $d = 1$ and choose $k^G(S,T)$ to be the linear kernel $k^G(S,T) = \langle S, T\rangle$ and k^{pos} to be any kernel on $\mathbb{R}^2$. In this case, $\widetilde{G}_1^2$ is identified as the unit circle S^1 on the plane and the inner product on $\widetilde{G}_1^2$ is inherited from the standard Euclidean metric on $\mathbb{R}^2$. Now let $\mu \doteq \delta_x \otimes \mathcal{H}^1|_{S^1}$ be the non-zero varifold defined as the product measure of a point mass in $\mathbb{R}^2$ and the uniform probability distribution on S^1. It's straightforward to verify that:

$$\|\mu\|_{W^*}^2 = k^{pos}(x,x)\int_0^{2\pi}\int_0^{2\pi}(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2)d\theta_1 d\theta_2 = 0.$$

Therefore, d_{W^*} is not a metric.

Recall that $\iota_W : W \hookrightarrow C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$ denotes the continuous embedding and $\iota_W^* : C_0(\mathbb{R}^n \times \widetilde{G}_d^n)^* \mapsto W^*$ denotes its adjoint operator. To recover a true distance on $\mathcal{V}_d$, one needs the previous map ι_W^* or equivalently the map

$$\mu \mapsto \int_{\mathbb{R}^n \times \widetilde{G}_d^n} k(\cdot,(y,T))d\mu(y,T),\ \mu \in C_0(\mathbb{R}^d \times \widetilde{G}_d^n)^* \tag{3.1}$$

to be injective. As follows from Theorem 6 in [49], this is in fact guaranteed when the kernel k on the product space $\mathbb{R}^n \times \widetilde{G}_d^n$ is C_0-universal, specifically

Theorem 11. *The pseudo-distance d_{W^*} induces a distance between signed measures of $\mathbb{R}^n \times \widetilde{G}_d^n$ if and only if k is C_0-universal on $\mathbb{R}^n \times \widetilde{G}_d^n$. In particular, a sufficient condition for d_{W^*} to be a distance on $\mathcal{V}_d$ is that k^{pos} and k^G are C_0-universal kernels on $\mathbb{R}^n$ and $\widetilde{G}_d^n$ respectively.*

Note that these conditions are more restrictive than in Theorem 10. To our knowledge, there is no simple characterization for general C_0-universal kernels on the

Grassmannian. However, within the assumption on the kernel in (1.28), one easily constructs C_0-universal kernels by restriction (based on the Plücker embedding) of C_0-universal kernels defined on the vector space $\Lambda^d(\mathbb{R}^n)$. For instance, we may choose k^G to be the oriented Gaussian kernel $k^G(S,T) \doteq \exp(-\frac{2(1-\langle S,T\rangle)}{\sigma^2})$, which comes from the restriction of the standard Gaussian kernel on $\Lambda^d(\mathbb{R}^n)$. This kernel is C_0-universal and leads to a metric on $\mathcal{V}_d$.

We now study more precisely the topology induced by the (pseudo) distance d_{W^*} on $\mathcal{V}_d$ in comparison with the classical ones defined in Section 1.3.3. First of all, we observe that, for any $\omega \in W$ with $\|\omega\|_W \leq 1$, one must have $\|\omega\|_\infty \leq c_W$, where c_W is the embedding constant of Proposition 4. Thus, for any μ and μ' in $\mathcal{V}_d$, we have

$$\|\mu - \mu'\|_{W^*} = \sup_{\omega \in W,\ \|\omega\|_W \leq 1} \int_{\mathbb{R}^d \times \widetilde{G}_d^n} \omega \ d(\mu - \mu') \leq c_W d_{op}(\mu, \mu'). \tag{3.2}$$

From the above inequalities we see that convergence in d_{op} implies convergence in d_{W^*}.

Remark 4. *With more assumptions on the regularity of the kernel k, namely if W is continuously embedded in $C_0^1(\mathbb{R}^d \times \widetilde{G}_d^n)$, following a similar reasoning as above, one obtains the bound $\|\mu - \mu'\|_{W^*} \leq c_W d_{BL}(\mu, \mu')$.*

Suppose μ_i converges to μ in narrow topology. Since the map $(\nu_1, \nu_2) \mapsto \nu_1 \otimes \nu_2$ is continuous with respect to the narrow topology (cf. Theorem 3.3 and 3.11 in [86]), we have

$$\begin{aligned}
\|\mu_i\|_{W^*}^2 &= \int_{(\mathbb{R}^d \times \widetilde{G}_d^n)^2} k((x,S),(y,T)) d\mu_i(x,S) d\mu_i(y,T) \\
&\to \int_{(\mathbb{R}^d \times \widetilde{G}_d^n)^2} k((x,S),(y,T)) d\mu(x,S) d\mu(y,T) \\
&= \|\mu\|_{W^*}^2,
\end{aligned}$$

as $i \to \infty$. Also, it is clear that $\lim_{i\to\infty} \langle \mu_i, \mu \rangle_{W^*} \to \|\mu\|_{W^*}^2$ and hence $\mu_i \to \mu$ with respect to d_{W^*}. To summarize the discussion above:

Proposition 7. *Let $\{\mu_i\}_i$ and μ be varifolds in $\mathcal{V}_d$ and assume that $\mu_i \to \mu$ with respect to the operator norm or the narrow topology, then $\mu_i \to \mu$ in W^*.*

Remark 5. *We emphasize that the result of Proposition 7 only requires the assumptions of Proposition 4 and thus holds whether ι is injective or not.*

As for the weak-* topology, with the C_0-universality assumption of Theorem 11 and restricting to varifolds with bounded total mass, we show that d_{W^*} induces a topology stronger than weak-* convergence:

Proposition 8. *Let $\mathcal{V}_{d,M} \doteq \{\mu \in \mathcal{V}_d \ s.t\ |\mu|(\mathbb{R}^n) \leq M\}$ for arbitrary fixed $M > 0$ and assume that k is C_0-universal. If a sequence $\{\mu_i\}_i$ converges to μ in $\mathcal{V}_{d,M}$ with respect to d_{W^*}, then $\mu_i \to \mu$ with respect to weak-* topology.*

Proof. Let $\{\mu_i\}_i$ and μ be varifolds in $\mathcal{V}_{d,M}$ and assume that $\lim_{i\to\infty} d_{W^*}(\mu_i, \mu) = 0$. For any $f \in C_0(\mathbb{R}^d \times \widetilde{G}_d^n)$ and $\varepsilon > 0$, there exists a $g \in W$ such that $\|g - f\| < \varepsilon/2M$. Then we obtain that $\mu_i \overset{*}{\rightharpoonup} \mu$ from the following inequalities:

$$|(\mu_i - \mu|f)| \leq |(\mu_i|f - g)| + |(\mu|g - f)| + |(\mu_i - \mu|g)| \leq \varepsilon + \|\mu_i - \mu\|_{W^*}\|g\|_W.$$

□

Note that the topology induced by d_{W^*} may be strictly finer on $\mathcal{V}_{d,M}$. Indeed, if $\rho(0), \gamma(1) > 0$, consider $\mu_i = \delta_{(x_i,S)}$, where $\lim_{i\to\infty} |x_i| = \infty$ and $S \in \widetilde{G}_d^n$ fixed. Then $\mu_i \overset{*}{\rightharpoonup} 0$ while $\|\mu_i\|_{W^*}^2 = \rho(0)\gamma(1) > 0$ for all i. Yet, by combining Propositions 3, 7 and 8, we have the following

Corollary 1. *Let $M > 0$ and $K \subset \mathbb{R}^n \times \widetilde{G}_d^n$ be a compact subset. If k is C_0-universal, then d_{W^*} metrizes the weak-* convergence of varifolds on $\mathcal{V}_{d,M,K} \doteq \{\mu \in \mathcal{V}_d \ s.t\ |\mu|(\mathbb{R}^n) \leq M,\ supp(\mu) \subset K\}$.*

In summary, C_0-universality provides a sufficient condition to obtain actual distances between varifolds that can be expressed based on the kernel function. Furthermore, the resulting topology is locally equivalent to the weak-* topology as well as the topology induced by the bounded Lipschitz distance. This equivalence will be of importance for the upcoming section.

3.2 Approximations by discrete varifolds

Having discussed the characterizations of the kernel metrics and their relationship with classical metrics on varifolds, we now move to investigate the problem of approximating general varifolds by discrete varifolds and investigate the consistency of the discretized registration problems (Theorem 13), which is the main result of this section.

3.2.1 Discrete approximations

In section 1.3.1, we have discussed how to approximate piecewise linear shapes given by meshes such as polygonal curves or triangulated surfaces using discrete varifolds. In the more general context of this work, a key issue is to construct similar discrete varifold approximations for more general and less structured objects. Specifically, given a varifold μ with finite total weight, can it be approximated by discrete varifolds and will approximations converge as $N \to +\infty$? This is the problem known as *quantization*, which has been studied intensively in the case of probability measures over Euclidean spaces [87] or manifolds [88], under specific regularity assumptions on those measures. In the situation of varifolds, an interesting recent work on this question is [89]. The authors prove that any rectifiable varifold with finite mass can be approximated by a sequence of discrete varifolds for the bounded Lipschitz distance and propose a numerical approach to approximate mean curvature measures based on discrete varifolds.

In this section, we first wish to extend approximation results to general oriented

varifolds of finite mass for both d_{BL} and d_{W^*} metrics.

Theorem 12. *Let*

$$\mathcal{V}_d^N := \left\{ \sum_{i=1}^N r_i \delta_{(x_i,T_i)} | r_i \in \mathbb{R}_+,\ x_i \in \mathbb{R}^n,\ T_i \in \widetilde{G}_d^n \right\}$$

be the (non-convex) cone of discrete varifolds with at most N Diracs. For any oriented varifold $\mu \in \mathcal{V}_d$ with $|\mu|(\mathbb{R}^n) < \infty$, there exists a sequence $\mu_N \in \mathcal{V}_d^N$ such that $\lim_{N\to\infty} d_{BL}(\mu_N, \mu) = 0$. Moreover, if μ has compact support, then we can assume that for all N, $\mathrm{supp}(\mu_\mathrm{N}) \subset \mathrm{K}$ *for some compact set $K \subset \mathbb{R}^n \times \widetilde{G}_d^n$ and*

$$d_{BL}(\mu_N, \mu) < \frac{C}{N^{1/(n+d(n-d))}},$$

where C is a constant that only depends on n, d and $\mathrm{supp}(\mu)$.

Proof. We first tackle the case of compactly supported μ. Without loss of generality, we may also assume that μ is a probability measure. Let $D = n + d(n-d)$ and $B \subset \mathbb{R}^n$ be a closed ball that contains $\mathrm{supp}|\mu|$. For brevity, we write $M \doteq B \times \widetilde{G}_d^n$. Since we can view M as a compact D-dimensional submanifold of $\mathbb{R}^n \times \Lambda^d(\mathbb{R}^n)$ (using Plücker embedding), M is also regular of dimension D (cf. [87]), i.e., $0 < \mathcal{H}^D(M) < \infty$ and there exist $c, r_0 > 0$, such that

$$\frac{1}{c} r^D \leq \mathcal{H}^D \llcorner M(B_r(a)) \leq c r^D,\ \forall a \in M,\ r \in (0, r_0).$$

Given $\varepsilon \in (0, 5r_0)$, by the 5-Times Covering Lemma (cf. [50]), these exists a subset $\mathcal{I} \subset M$, such that $M \subset \cup_{x\in\mathcal{I}} B_\varepsilon(x)$ and $B_{\varepsilon/5}(x) \cap B_{\varepsilon/5}(y) = \varnothing$ for all $x \neq y \in \mathcal{I}$. Therefore,

$$\mathcal{H}^D(M) \geq \sum_{x\in\mathcal{I}} \mathcal{H}^D(M \cap B_{\varepsilon/5}(x)) \geq \frac{|\mathcal{I}|\varepsilon^D}{c5^D}.$$

We can thus obtain a partition $\{A_i\}_{i=1,\cdots,|\mathcal{I}|}$ of M from the the collection $\{B_\varepsilon(x) \cap M\}_{x\in\mathcal{I}}$ which satisfies $\sup_i \mathrm{diam}(A_i) < \varepsilon$ and

$$|\mathcal{I}| \leq \frac{c5^D \mathcal{H}^D(M)}{\varepsilon^D}.$$

Let $r_i = \mu(A_i)$ and $(x_i, T_i) \in A_i$ and define $\nu = \sum_{i=1}^{|\mathcal{I}|} r_i \delta_{(x_i,T_i)}$. For any $\varphi \in \mathrm{Lip}_1(\mathbb{R}^n \times \widetilde{G}_d^n)$, with $\|\varphi\|_\infty \leq 1$, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n\times\widetilde{G}_d^n} \varphi(x,T)d\nu - \int_{\mathbb{R}^n\times\widetilde{G}_d^n} \varphi(x,T)d\mu\right| &= \left|\sum_{i=1}^{|\mathcal{I}|}\left(\mu(A_i)\varphi(x_i,T_i) - \int_{A_i}\varphi(x,T)d\mu\right)\right| \\ &\leq \sum_{i=1}^{|\mathcal{I}|}\int_{A_i}|\varphi(x_i,T_i)-\varphi(x,T)|d\mu \\ &< \sum_{i=1}^{|\mathcal{I}|}\varepsilon\mu(A_i) = \varepsilon.\end{aligned}$$

Taking the supremum over all $\omega \in \mathrm{Lip}_1(\mathbb{R}^n \times \widetilde{G}_d^n)$ with $\|\varphi\|_\infty \leq 1$, we obtain $d_{BL}(\mu,\nu) < \varepsilon$. Then for each $N \in \mathbb{N}$, we can choose $\varepsilon_N = 5(C\mathcal{H}^D(M)/N)^{1/D}$ and we obtain $\mu_N \in \mathcal{V}_d^N$ such that

$$d_{BL}(\mu,\mu_N) < \frac{5C^{1/D}(\mathcal{H}^D(M))^{1/D}}{N^{1/D}}$$

and in particular $\lim_{N\to+\infty} d_{BL}(\mu,\mu_N) = 0$.

Suppose now that $\mathrm{supp}(\mu)$ is not compact: we show that for any $\varepsilon > 0$, there exists a discrete varifold ν such that $d_{BL}(\mu,\nu) < \varepsilon$. Choose a compact set $K \subset \mathbb{R}^n \times \widetilde{G}_d^n$ such that $\mu(\mathbb{R}^n \times \widetilde{G}_d^n \setminus K) < \varepsilon/2$. From the previous case, we can find a discrete varifold ν such that $d_{BL}(\mu \llcorner K, \nu) < \varepsilon/2$, and hence $d_{BL}(\mu,\nu) < \varepsilon$. □

Note that the proposition clearly holds for non-oriented varifolds as well. Another direct consequence, thanks to proposition 7 and remark 4, is the following corresponding statement for d_{W^*}:

Corollary 2. *With the assumptions from proposition 4, one also has* $\lim_{N\to\infty} d_{W^*}(\mu_N,\mu) = 0$. *If in addition* $W \hookrightarrow C_0^1(\mathbb{R}^n \times \tilde{G}_d^n)$, *an equivalent upper bound as in Theorem 12 holds for* $d_{W^*}(\mu,\mu_N)$.

We should point out that the asymptotic convergence rate given by the previous upper bound is rather slow, especially as the dimensions d and n grow. This is

however under very mild assumptions on the varifold μ. We expect much better convergence properties for certain specific classes of varifolds, for instance assuming Alfors regularity as in [88], although we leave such questions for future investigation.

3.2.2 Optimal approximating sequence

In addition to the asymptotic approximation results of the previous section, we now want to construct such sequences of discrete approximating varifolds. Given any $\mu \in \mathcal{V}_d$ and $N \in \mathbb{N}$, a natural idea is to look for the optimal discrete varifold in $\mathcal{V}_d^N$ that approximates μ in terms of the metric d_{W^*}. Due to the intricate structure of the set $\mathcal{V}_d^N$ (infinite-dimensional non-convex cone), this is far from a straightforward problem. Several different approaches in some simpler contexts have been proposed to circumvent this issue, which we briefly recap. One possibility is to restrict to finite-dimensional vector spaces of $\mathcal{V}_d$ (e.g. generated by finite sets of Diracs). Works such as [90–92] for instance, which are focused on the model of currents, consider dictionaries of Diracs defined on a predefined grid of point positions in $\mathbb{R}^n$. Then the problem can be recast as the one of finding sparse approximations of μ in such a dictionary. It remains a NP hard problem but solutions can be approached either through greedy algorithms like *orthogonal matching pursuit* as proposed in [90, 91] or by considering the L^1 relaxation formulation leading to a standard convex LASSO program. Such ideas apply well to the specific situation of currents mainly as a result of the inherent linearity of this model: indeed, at any iteration of a matching pursuit procedure, once the optimal position of a Dirac is found, the corresponding direction vector and weight are explicitly determined. This allows to limit the search over grid of points in the spatial domain only. Unfortunately, for the general oriented varifold metrics we consider in this paper, such a property no longer holds and, as a result, these methods would involve very large dictionaries defined on grids on the product $\mathbb{R}^n \times \widetilde{G}_d(\mathbb{R}^n)$. Such an increase in dimension makes the approach numerically

impractical as soon as $n \geq 3$ and $d \geq 2$. Another downside is that the use of finite dictionaries and greedy algorithms like matching pursuit is not guaranteed to give an optimal approximation of varifolds for a given number N of Diracs. The approach we develop in this section consists instead in directly tackling the non-convex problem of computing the projection onto $\mathcal{V}_d^N$ for the class of kernel metrics d_{W^*}. It shares some connection with the recent work of [93] that considers a related problem for standard measures defined on the torus $\mathbb{R}^n/\mathbb{Z}^n$.

Fix a varifold $\mu_* \in \mathcal{V}_d$. For any $N \in \mathbb{N},\ N \geq 1$, we seek $\mu_N \in \mathcal{V}_d^N$ that is closest to μ_* for d_{W^*}, namely

$$\mu_N = \operatorname*{argmin}_{\mu \in \mathcal{V}_d^N} \|\mu - \mu_*\|_{W^*} \tag{3.3}$$

By construction, if $|\mu|(\mathbb{R}^n) < \infty$ then Corollary 2 will imply that (μ_N) converges to μ in the metric d_{W^*}. We only need to ensure that such a projection is well defined, which is the object of the following proposition:

Proposition 9. *Suppose all assumptions in proposition 4 hold and the separable kernel has the form as in (1.28). We further assume that the functions ρ and γ defining the kernels are non-negative. Then for any $\mu \in \mathcal{V}_d$ and $N \in \mathbb{N}$, there exists $\mu_N \in \mathcal{V}_d^N$ such that $\mu_N = \operatorname{argmin}_{\nu \in \mathcal{V}_d^N} \|\mu - \mu_*\|_{W^*}$*

Proof. Let $\mu_m = \sum_{i=1}^N r_i^m \delta_{(x_i^m, T_i^m)}$ be a minimizing sequence, $K_W : W^* \mapsto W$ be the dual operator and $f := K_W(\mu)$. Without loss of generality, we may assume that there is a $N_1 \leq N$ such that $\sup_{1 \leq i \leq N_1} |x_i^m|$ remains bounded and $\inf_{M+1 \leq i \leq N} |x_i^m|$ tends to ∞ as $m \to \infty$.

Observe that $\sup_{1 \leq i \leq N} \{r_i^m\}$ must be bounded. If it's not bounded, then from the assumptions that $\rho, \gamma \geq 0$, we obtain

$$\begin{aligned}\|\mu_m - \mu_*\|_{W^*} &\geq \|\mu_m\|_{W^*} - \|\mu_*\|_{W^*} \\ &= \sqrt{\sum_{i,j=1}^{N} r_i^m r_j^m \rho(|x_i^m - x_j^m|^2)\gamma(\langle T_i^m, T_j^m\rangle)} - \|\mu_*\|_{W^*} \\ &\geq \sqrt{\sum_{i=1}^{N} (r_i^m)^2 \rho(0)\gamma(1)} - \|\mu_*\|_{W^*} \to \infty\end{aligned}$$

as $m \to \infty$, which is absurd.

Since r_i^m, T_i^m, $f(x_i^m, T_i^m)$ and

$$A_m := (\rho(|x_i^m - x_j^m|^2)\gamma(\langle T_i^m, T_j^m\rangle))_{1\leq i,j\leq N}$$

are all bounded sequences of m, we may replace them by convergent subsequences, thus we could assume that

$$\lim_{m\to\infty} r_i^m = r_i, \quad \lim_{m\to\infty} T_i^m = T_i, \ \lim_{m\to\infty} f(x_i^m, T_i^m) = f_i, \quad \lim_{m\to\infty} A_m = A.$$

Since $\rho \in C_0(\mathbb{R})$, the matrix A must has the following form:

$$A = \begin{pmatrix} B_1 & \mathbf{0} \\ \mathbf{0} & B_2 \end{pmatrix},$$

where B_1 and B_2 are N_1-by-N_1 and $N - N_1$-by-$N - N_1$ semi-positive definite matrices. Combining this with the assumption $f \in C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$, we obtain

$$\lim_{m\to\infty} \|\mu_m - \mu_*\|_{W^*}^2 = \boldsymbol{r'}^T B_1 \boldsymbol{r'} + \boldsymbol{r''}^T B_2 \boldsymbol{r''} - 2\boldsymbol{f'}^T \boldsymbol{r'} + \|\mu_*\|_{W^*}^2,$$

where $\boldsymbol{r'} = (r_1, \cdots, r_{N_1})$, $\boldsymbol{r''} = (r_{N_1+1}, \cdots, r_N)$, and $\boldsymbol{f'} = (f_1, \cdots, f_{N_1})$. Since $\sup_{1\leq i\leq N_1} |x_i^m|$ is bounded we can assume that $\lim_{m\to\infty} x_i^m = x_i$, $1 \leq i \leq N_1$. Let $\mu := \sum_{i=1}^{N_1} r_i \delta_{(x_i,u_i)}$, then

$$\|\mu - \mu_*\|_{W^*}^2 = \boldsymbol{r'}^T B_1 \boldsymbol{r'} - 2\boldsymbol{f'}^T \boldsymbol{r'} + \|\mu_*\|_{W^*}^2 \leq \lim_{m\to\infty} \|\mu_m - \mu_*\|_{W^*}^2.$$

Hence μ is a minimizer. □

However, in general this projection is not unique. We also point out that the existence is a priori not guaranteed if kernels ρ and γ take negative values. It is so far an open question to determine to what extent one could generalize the result of Proposition 9, one particular but important case being the one of current metrics obtained for $\gamma(t) = t$ which is not covered by our result.

As written in the proof of Proposition 9, (3.3) is equivalent to the optimization problem:

$$(r_i, x_i, T_i) = \underset{(w_i, y_i, S_i)}{\operatorname{argmin}} \| \sum_{i=1}^{N} w_i \delta_{(y_i, S_i)} - \mu_* \|_{W^*}^2 \tag{3.4}$$

Remark 6. *Any solution must satisfy first order optimality conditions obtained by differentiating $\|\mu_N - \mu\|_{W^*}$ with respect to the (r_k, x_k, T_k). In particular, we have*

$$\begin{aligned} 0 &= \frac{\partial \|\mu_N - \mu_*\|_{W^*}^2}{\partial r_k} \\ &= 2\Bigg(\sum_{i=1}^{N} r_i \rho(|x_i - x_k|^2) \gamma(\langle T_i, T_k \rangle) - \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \rho(|x_k - x|^2) \gamma(\langle T_k, T \rangle) d\mu_*(x, T) \Bigg). \end{aligned}$$

which gives $\langle \mu_N - \mu_*, \mu_N \rangle_{W^*} = 0$. It shows that for any $N \in \mathbb{N}$, $\|\mu_N\|_{W^*} \leq \|\mu_*\|_{W^*}$.

On the numerical side, we solve the projection problem (3.3) with the frame representation introduced in section 2.2.1. In this setting, the triple $(r_i, x_i, T_i)_{i=1,\cdots,N}$ is identified with the state $q = (x_i, u_i^{(1)}, \cdots, u_i^d)_{i=1,\cdots,N}$, where $T_i = u_i^{(1)} \wedge \cdots \wedge u_i^{(d)} / |u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}|$ and $r_i = |u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}|$. Then the solution to the projection problem (3.3) can be computed by an iterative descent strategy on the vector $q = (x_i, u_i^{(1)}, \cdots, u_i^{(d)})_{i=1,\cdots,N}$. The gradient of

$$\begin{aligned} q \mapsto & \|\mu^q - \mu_*\|_{W^*}^2 \\ &= \sum_{i,j=1}^{N} |u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}| |u_j^{(1)} \wedge \cdots \wedge u_j^{(d)}| \rho(|x_i - x_j|^2) \gamma(\langle T_i, T_j \rangle) \\ &- 2 \sum_{i=1}^{N} \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \rho(|x_i - x|^2) \gamma(\langle T_i, T \rangle) d\mu_*(x, T) \\ &+ \int_{(\mathbb{R}^n \times \widetilde{G}_d^n)^2} \rho(|x - y|^2) \gamma(\langle T, S \rangle) d\mu_*(x, T) d\mu_*(y, S), \end{aligned}$$

with respect to the x_i and $u_i^{(l)}$. The optimization itself is done using a limited memory BFGS algorithm from the HANSO library [81] which we typically initialize by taking a random subset of N Diracs composing the varifold μ_*. Note that one of the main downside of this projection algorithm, in contrast with the previously mentioned approach of fixing a dictionary and solving a convex sparse decomposition problem, is that we can provide no general guarantees of convergence to a global minimum of (3.3). Results of this algorithm are discussed below in Section 3.3.

3.2.3 Γ-convergence of registration functionals

Ultimately, our purpose is to use the previous approximating discrete varifolds μ_N to approximate the diffeomorphic registration problem (2.1). The natural question that arises is whether replacing the source varifold μ_0 by its projections μ_N in (2.1) still leads to reasonable approximations (at least asymptotically) of optimal deformation fields for the original problem. In this section, we address this by showing a Γ-convergence property for these variational problems. We point out that our setting and the following proof differ quite a bit from previous results of the same type that were dealing with the specific case of surface triangulations such as [94] or [95].

To obtain such convergence results for solutions of variational problems, one usually requires the approximating sequence to possess certain nice properties. Specifically, assuming $\mu \in \mathcal{V}_d$ with compact support and finite mass and $\{\mu_N\} \subset \mathcal{V}_d^N$ such that $\lim_{N\to\infty} \|\mu_N - \mu\|_{W^*} = 0$, we will need that $\bigcup_N \operatorname{supp}(\mu_N) \subset K$ for some compact set $K \subset \mathbb{R}^n$ or that $\sup_N |\mu_N|(\mathbb{R}^n) < \infty$. Unfortunately, this does not hold in general since convergence in d_{W^*} does not allow to control the support nor the total mass of the sequence μ_N.

Yet, provided that $\bigcup_N \operatorname{supp}(\mu_N) \subset K$, we can actually retrieve the boundedness of the total mass. We assume in what follows that the kernels are such that $\rho(0) > 0$ and $\gamma(1) > 0$.

Lemma 2. *Let $\{\mu_N\}$ be a sequence of discrete varifolds with finite mass such that there exists a compact $K \subset \mathbb{R}^n$ with* $\mathrm{supp}(|\mu_N|) \subset K$ *for all N. We assume that $\{\|\mu_N\|_{W^*}\}$ is bounded. Then $\{|\mu_N|(\mathbb{R}^n)\}$ is bounded.*

Proof. We prove it by contradiction. Assume that $(|\mu_N|(\mathbb{R}^n))_{N\geq 1}$ is unbounded. Then, up to extracting a subsequence, we can assume that $|\mu_N|(\mathbb{R}^n) \to +\infty$. Let's write $\mu_N = \sum_{i=1}^{p_N} r_{i,N}\delta_{(x_{i,N},T_{i,N})}$. Thus $|\mu_N|(\mathbb{R}^n) = \sum_{i=1}^{p_N} r_{i,N} \to +\infty$. Since ρ and γ are continuous and $\rho(0), \gamma(1) > 0$, we can find compact subsets $A \subset K$ and $B \subset \tilde{G}_d^n$ with diameters small enough, so that: $\inf_{x,y\in A} \rho(|x-y|^2) > m > 0$, $\inf_{u,v\in B} \gamma(\langle u, v\rangle) > m' > 0$ and $\lim_{N\to\infty} \sum_{i\in\mathcal{I}_N} r_{i,N} = \infty$, where $\mathcal{I}_N := \{i : (x_{i,N}, u_{i,N}) \in A \times B\}$. It follows that, as $N \to \infty$,

$$\begin{aligned}
\|\mu_N\|_{W^*}^2 &= \sum_{i=1}^{p_N}\sum_{j=1}^{p_N} r_{i,N}r_{j,N}\rho(|x_{i,N}-x_{j,N}|^2)\gamma(\langle T_{i,N}, T_{j,N}\rangle) \\
&\geq mm' \sum_{i,j\in\mathcal{I}_N} r_{i,N}r_{j,N} \\
&= mm' \left(\sum_{i\in\mathcal{I}_N} r_{i,N}\right)^2 \to \infty,
\end{aligned}$$

which is a contradiction. □

Lemma 2 suggests that one should enforce the uniform compactness of the supports of the μ_N. To do so in the context of the projection approach of the previous sections, we consider solving the optimization problem (3.3) with the additional constraint that the support of μ_N stays in a compact set containing $\mathrm{supp}(|\mu|)$. We still have to verify the convergence of the resulting sequence:

Lemma 3. *Let μ_0 be a varifold with finite mass and K be a compact set in $\mathbb{R}^n$ which contains* $\mathrm{supp}(|\mu_0|)$. *Construct the approximating sequence of μ_0 by solving the following constrained optimization problem:*

$$\begin{aligned}
&\mu_{K,N} = \operatorname*{argmin}_{\nu\in\mathcal{V}_d^N} \|\nu - \mu\|_{W^*} \\
&\textit{subject to } \mathrm{supp}(|\nu|) \subset K.
\end{aligned}$$

Then $\mu_{K,N}$ converges to μ_0 in d_{W^*} and, if the kernel k is C_0-universal, it also converges in d_{BL}.

Proof. Thanks to Theorem 12 and Lemma 2, we immediately get that $\|\mu_{K,N} - \mu_0\|_{W^*} \to 0$ as $N \to \infty$ and $\sup_N(|\mu_{K,N}|(\mathbb{R}^n)) < \infty$. Moreover, if k is C_0-universal, then by Proposition 8 it implies that $\mu_{K,N} \overset{*}{\rightharpoonup} \mu_0$. Since $\bigcup_N \mathrm{supp}(|\mu_{\mathrm{K,N}}|) \subset \mathrm{K}$ and $\sup_N(|\mu_{K,N}|(\mathbb{R}^n)) < \infty$, weak-* convergence implies that $\mu_{K,N}$ converges to μ_0 in d_{BL} by Proposition 3. □

We now get to the main result of this section. We assume that the source/template varifold μ_0 is compactly supported and we fix K is a compact subset of $\mathbb{R}^n$ that contains $\mathrm{supp}(|\mu_0|)$. Then for any $N \in \mathbb{N},\ N \geq 1$, $\mu_{K,N}$ is defined as in Proposition 3 and we introduce the following energy functionals $E_N :\ L^2([0,1], V) \to \mathbb{R}_+$:

$$\begin{aligned} &E_N(v) \doteq \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \lambda \|\mu_{K,N}(1) - \mu_{tar}\|_{W^*}^2 \\ &\textit{subj to } \begin{cases} \partial_t \varphi_t^v = v_t \circ \varphi_t^v,\ \varphi_0^v = id \\ \mu_{K,N}(t) = (\varphi_t^v)_{\#}\mu_{K,N} \end{cases} \end{aligned} \tag{3.5}$$

which are the equivalent to the energy E of the original problem (2.1) but replacing the template varifold μ_0 by its approximations $\mu_{K,N}$.

Theorem 13. *With the above notations, we assume that the reproducing kernel k of W is C_0-universal and satisfies all the conditions of Proposition 9. We also assume the continuous embedding $V \hookrightarrow C_0^2(\mathbb{R}^n, \mathbb{R}^n)$. Then, the sequence of functionals E_N Γ-converges to E for the weak topology on $L^2([0,1], V)$. Consequently, if v_N is a global minimizer of E_N for each $N \geq 1$, then (v^N) is bounded in $L^2([0,1], V)$ and every cluster point for the weak topology of $L^2([0,1], V)$ is a global minimum of E.*

Proof. We first show that whenever v^N converges to $\bar{v}$ weakly in $L^2([0,1], V)$, we have

$$E(\bar{v}) \leq \liminf_{N \to \infty} E_N(v^N).$$

Since $v \mapsto \int_0^1 \|v\|_V^2 dt$ is lower semicontinuous with respect to the weak topology, we only need to prove the following,

$$\lim_{N\to\infty} \|(\varphi_1^{v^N})_{\#}\mu_{K,N} - \varphi_1^{\bar{v}} \cdot \mu_0\|_{W^*} = 0. \tag{3.6}$$

For all $\omega \in W$ with $\|\omega\|_W \leq 1$, we have

$$\begin{aligned}
\left|\left((\varphi_1^{v^N})_{\#}\mu_{K,N} - (\varphi_1^{v^N})_{\#}\mu_0|\omega\right)\right| &= \left| \int_{K\times\tilde{G}_d^n} J_S\varphi_1^{v^N}(x)\omega(\varphi_1^{v^N}(x), d_x\varphi_1^{v^N}\cdot S)d(\mu_{K,N}-\mu_0)\right| \\
&\leq C_1 \int_{K\times\tilde{G}_d^n} \sup_{N\geq 1} J_S\varphi_1^{v^N} d(\mu_{K,N}-\mu_0) \\
&\leq C_1 \int_{\mathbb{R}^n\times\tilde{G}_d^n} g(x,T)d(\mu_{K,N}-\mu_0),
\end{aligned}$$

where $g \in C_c(\mathbb{R}^n \times \tilde{G}_d^n)$ and $\sup_{N\geq 1} J_S\varphi_1^{v^N} \leq g(x,S)$, for all $(x,S) \in K \times \tilde{G}_d^n$. Similar to the computation done in the proof of Theorem 9, we see that

$$\left|\left((\varphi_1^{v^N})_{\#}\mu_0 - (\varphi_1^{\bar{v}})_{\#}\mu_0|\omega\right)\right| \leq C_2\|(\varphi_1^{v^N} - \varphi_1^{\bar{v}})|_K\|_{1,\infty}.$$

Taking supremum over all $\omega \in W$ with $\|\omega\|_W \leq 1$, we obtain the following inequality,

$$\|(\varphi_1^{v^N})_{\#}\mu_{K,N} - (\varphi_1^{\bar{v}})_{\#}\mu_0\|_{W^*} \leq C_1(\mu_{K,N}-\mu_0|g) + C_2\|(\varphi_1^{v^N} - \varphi_1^{\bar{v}})|_K\|_{1,\infty}.$$

From Proposition 3, $\mu_{K,N}$ converges to μ_0 in the narrow topology. Hence the right hand side in the equation above converges to 0 as $N \to \infty$. This proves (3.6).

Second, we need to show that for each $\bar{v} \in L^2([0,1],V)$, there exists a sequence v^N converging to $\bar{v}$ weakly such that

$$E(\bar{v}) \geq \limsup_{N\to\infty} E_N(v^N).$$

In fact, it suffices here to take v^N to be the constant sequence $v^N = \bar{v}$ since, by a similar argument to the proof of (3.6), it leads to

$$\lim_{N\to\infty} \|(\varphi_1^{\bar{v}})_{\#}\mu^N - (\varphi_1^{\bar{v}})_{\#}\cdot\mu\|_{W^*} = 0 \tag{3.7}$$

and thus implies that

$$\limsup_{N\to\infty} E_N(v^N) = \lim_{N\to\infty} E_N(\bar{v}) = E(\bar{v}).$$

□

Note that we stated the result of Theorem 13 in the situation where only the source varifold μ_0 is approximated by the projection approach that we presented in the previous sections but one can easily extend it to the scenario in which both source and target are replaced by discrete approximating sequences, the conclusion being the same in that case.

3.3 Numerical results

In this last section, we illustrate the previous mathematical results by showing a few numerical examples of the varifold quantization procedure introduced in Section 3.2, and its interplay with the registration algorithm. Specifically, we wish to numerically validate the statements of Corollary 2 and Theorem 13. We shall consider the following protocol. Starting from a highly sampled shape (that we treat as the groundtruth) for which the associated varifold μ_0 is composed of a very high number of Diracs, we compute the compressed varifolds given by the μ_N of (3.3) for increasing values of N and evaluate the resulting quantization error in terms of the d_{W^*} metric. Then we solve the registration problems to a fixed target μ_{tar} from the source varifolds given by the μ_N in lieu of μ_0, and compare the estimated solutions to the registration of the groundtruth. For comparison, we will evaluate the total energy $E(v^N)$ of the estimated deformation fields v^N for the original problem, i.e.

$$E(v^N) = \int_0^1 \|v_t^N\|_V^2 dt + \lambda \|(\varphi_1^{v^N})_{\#}\mu_0 - \mu_{tar}\|_{W^*}^2.$$

We shall also compare this overall approach against the alternative idea of directly subsampling the original meshes and registering those subsampled shapes with point

set mesh LDDMM i.e. with the approach of [45] and [55].

We begin with a 1-varifold toy example (Figure 3-2) given by the curves shown in Fig. 3-1 from the Kimia database. These very simple curves segmented from binary images have a relatively high number of points to start with (368 vertices and edges). We look first at how well they can be approximated with smaller number of Diracs through the quantization approach described above. The upper row shows the plot of the relative approximation error $\|\mu_N - \mu_0\|_{W^*}/\|\mu_0\|_{W^*}$ of the source curve as a function of N (blue) as well as the same error in varifold norm when instead the curve is uniformly subsampled (green). Consistent with the fact that varifold quantization should provide the optimal error rate at a given N, we observe that the error is indeed smaller than with the subsampling approach. We also display a few of the quantized μ_N for several values of N. As a second step, we compute the optimal deformations from the reduced shapes to the fixed target and compare their registration energies to the "groundtruth" $E(v^*)$ estimated from the full resolution source shape. The corresponding plots for the quantization versus subsampling methods are shown on the lower row in blue and green respectively. It suggests again a faster convergence to the optimal energy $E(v^*)$ with the quantization strategy, although the difference between the two methods is rather tenuous in this example.

Those effects can be much more significant in the two-dimensional case. We emphasize it with the triangulated heart surfaces of Fig. 3-2 (data courtesy of C. Chnafa, S. Mendez and F. Nicoud, University of Montpellier). The source surface has a total of 42448 triangles leading to the same number of Diracs for the source 2-varifold μ_0 and thus compressing the representation may be in that case quite critical from a computational standpoint. Indeed, computing the groundtruth matching at full resolution takes more than 7 hours (68s per iteration) in this case. We again compare two approaches: our quantization algorithm applied to μ_0 versus directly subsampling the triangulated surface itself (we use here the reducepatch function

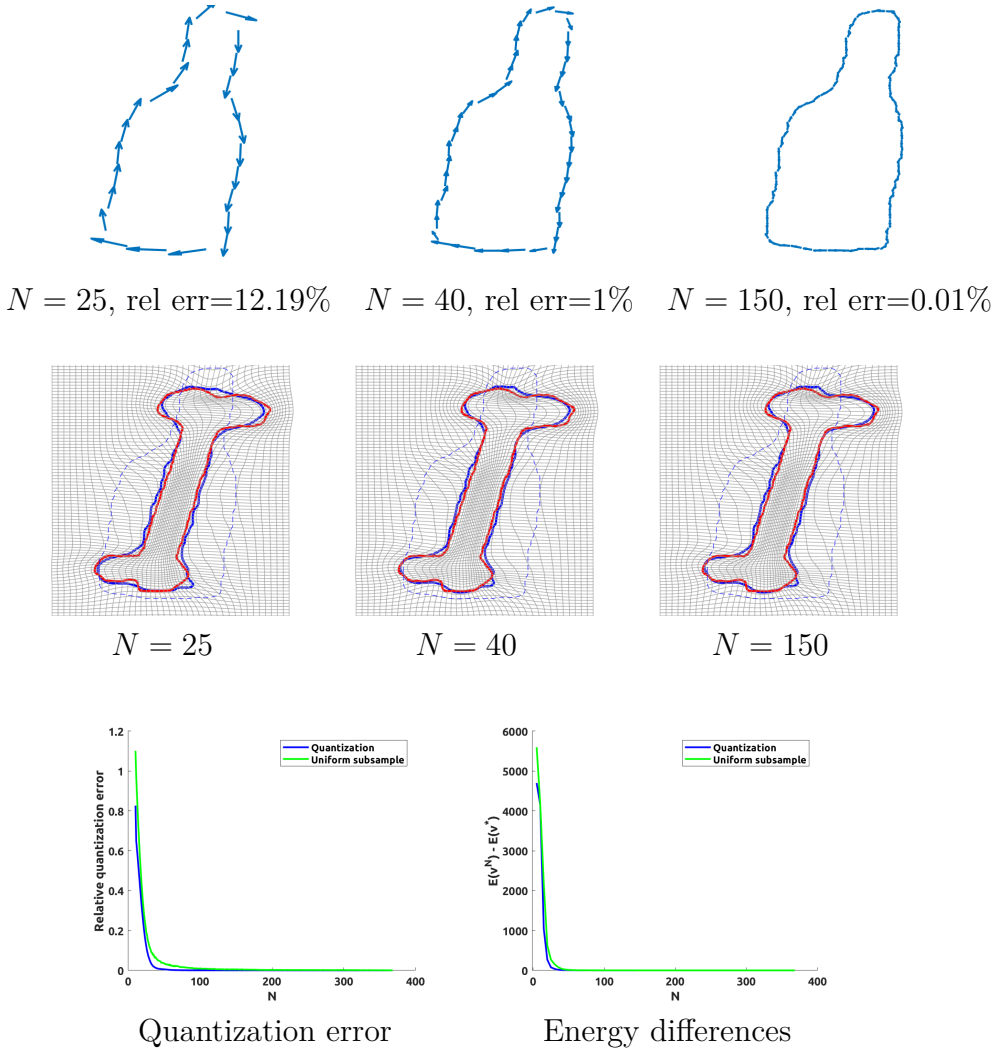
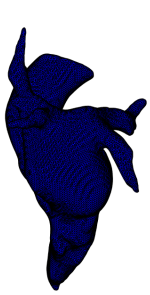
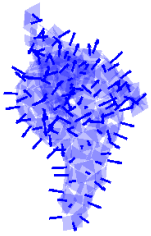
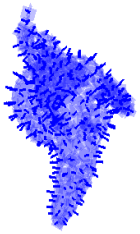


Figure 3-1. Compression and registration of 1-varifolds. The first row shows the results of the quantization algorithm on the 1-varifold associated to the source shape for different values of N. The second row shows the registration results using the apprroximated source in the first row. The plot on the left of the third row shows the relative quantization errors of the quantization (blue curve) and the errors obtained with a uniform subsampling scheme (green curve). The plot on the right of the third row shows the difference to the groundtruth optimal energy when solving the registration problem from the approximate source given by the varifold quantization (blue) and the direct subsampling approach (green).

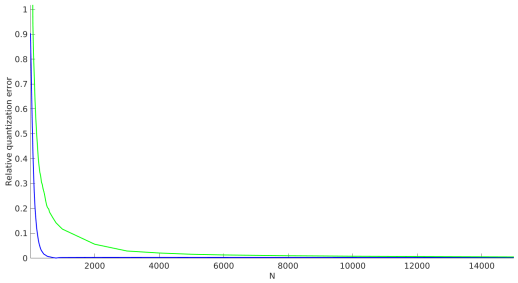


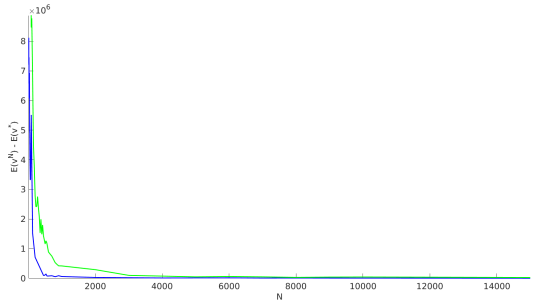
Source surface (42448 triangles) Target surface (50352 triangles)





$N = 65$, rel err=28% $N = 125$, rel err=7.1% $N = 375$, rel err=0.07%





Relative quantization error plot Energy differences: $E(v^n) - E(v^*)$.

Figure 3-2. Compression and registration of 2-varifolds. On top, the source and target triangulated surfaces. The second row shows the results of the quantization algorithm on the 2-varifold associated to the source shape for different values of N.The left plot on the third row shows the relative quantization errors (blue curve) and compared to the errors obtained with a mesh subsampling scheme (green curve). The right plot on the third row shows the difference to the groundtruth optimal energy when solving the registration problem from the approximate source given by the varifold quantization (blue) and the mesh subsampling approach (green).

in MATLAB to reduce the initial mesh to a given number of triangles). For both methods, we compute the relative approximation error $\|\mu_N - \mu_0\|_{W^*}/\|\mu_0\|_{W^*}$ with different values of N, the number of Diracs (resp. triangles) of the compressed varifold (resp. mesh). This is shown on the left second row in Fig. 3-2. Unsurprisingly, we see that the quantization approach leads to a much faster decrease in the error as a function of N but that in addition we obtain a very good approximation of μ_0 with only a small fraction of the initial number of Diracs. Some of the quantized varifolds μ_N are displayed in the figure. We also evaluate how well the solution of the registration problem to the target varifold or surface can be approximated based on the quantized source shapes. With v^* being a numerical solution for the groundtruth and v^N the solutions based on the quantized source shapes, the third row of Fig. 3-2 shows the difference of the energies $E(v^N) - E(v^*)$. We observe again a faster convergence towards the groundtruth optimal energy with the varifold quantization than with mesh subsampling.

Chapter 4

Diffeomorphic registration with density changes for the analysis of imbalanced shapes

In previous chapters, we have presented registration of varifold representation based on pure diffeomorphic transformations. However, there are important limitations to this approach when dealing with registration of what we term generically as *imbalanced shapes*, namely in the situation where the representing measures display significant variations of mass or density. A motivating example is the case of white matter fiber bundles in which one can expect not only variations in the overall geometry of the bundle but also changes in the number (i.e. density) of fiber curves in each bundle. Diffeomorphic registration of fiber bundles [92] thus typically rely on an ad hoc renormalization/simplification step to compensate for fiber density inconsistencies. Another quite common situation is when a shape, for instance an anatomical surface, is only partially or sparsely known due to acquisition or segmentation issues. We propose to take further advantage of the flexibility of the measure setting established in the previous chapters by augmenting the diffeomorphic component of LDDMM with a global or local change of density of the source measure to account for potential mass imbalance. We then introduce a generalized registration models in which deformation and density change are estimated jointly. Our models differ from the metamorphosis

setting of [96] in that we consider a more general class of measures better adapted to curves and surfaces but also restrict to transformations of the density only, thus avoiding the singularity issues described in [96]. Results of this chapter were partially published in [97].

4.1 Global weight change model

Although the pushforward action of diffeomorphisms on $\mathcal{V}_d$ defined in section 1.3.2 does allow to transform the mass of measures through the Jacobian determinant of the deformation, registering measures with important inconsistencies or density variations may lead to unnatural or even degenerate optimal deformations as we will show in some of the examples of Section 4.3. Our goal is thus to augment (1.19) with a complementary process to simultaneously modify the density of the measure. In this section, we will focus on a global model with a single common density rescaling factor and we will introduce an extension of the approach to deal with local changes in density in section 4.2.

4.1.1 Augmented optimal control problem

Adopting the notations of the previous chapters, we introduce a complementary rescaling factor $\alpha \in \mathbb{R}^+$, which is a nonnegative number acting as a global multiplicative factor on the measure μ:

$$(\alpha.\mu|\omega) \doteq \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha\omega(x,T)d\mu(x,T).$$

Under this extended setting, we formulate the registration of a source μ_0 to a target μ_{tar} as the following new optimization problem:

$$\min_{v,\alpha} E(v,\alpha) \doteq \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \frac{\tau}{2}(\alpha-1)^2 + \lambda\|\alpha.\mu(1) - \mu_{tar}\|_{W^*}^2 \tag{4.1}$$

subject to $\mu(t) \doteq (\varphi_t^v) \cdot \mu_0$. The rescaling factor α is here penalized by the simple squared difference with $\alpha = 1$ weighted by a fixed coefficient $\tau > 0$ and one can see formally that letting $\tau \to +\infty$ imposes $\alpha = 1$ and (4.1) then reduces to the previous LDDMM registration problem (2.1). It is relatively easy to show that the existence of solutions continue to hold for this more general problem:

Proposition 10. *Assume that V is continuously embedded in $C_0^2(\mathbb{R}^n, \mathbb{R}^n)$, W is continuously embedded in $C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$,* $\mathrm{supp}(\mu_0) \subset \mathrm{K}$ *for some compact subset K of $\mathbb{R}^n \times \widetilde{G}_d^n$ and that $\tau > 0$. Then there exist $(v, \alpha) \in L^2([0,1], V) \times \mathbb{R}_+$ achieving the infimum in (4.1).*

Proof. The proof follows roughly the same arguments as for Theorem 9. Let us consider a minimizing sequence (v^j, η^j) for problem (4.1). Here again the sequence (v^j) is bounded in $L^2([0,1], V)$ so we can assume (by extracting a subsequence if necessary) that (v^j) converges weakly to some $\bar{v}$ in $L^2([0,1], V)$ which implies that $\|(\varphi_1^{v_j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} \to 0$ as $j \to +\infty$. On the other hand, due to the second term in the functional (and since $\tau > 0$) we also have that (α^j) is a bounded sequence of $\mathbb{R}_+$ and thus up to extraction of another sequence, we may assume that $\alpha^j \to \bar{\alpha}$ for some $\bar{\alpha} \in \mathbb{R}_+$.

We then have $(\alpha^j - 1)^2 \to (\bar{\alpha} - 1)^2$ and from the lower semicontinuity of the deformation energy:

$$\int_0^1 \|\bar{v}\|_V^2 dt \leq \liminf_{j \to +\infty} \int_0^1 \|v^j\|_V^2 dt$$

In addition, for any $\omega \in W$:

$$\begin{aligned}
&\left|\left(\alpha^j(\varphi_1^{v^j})_\# \mu_0 - \bar{\alpha}(\varphi_1^{\bar{v}})_\# \mu_0 | \omega\right)\right| \\
&= \left| \int_K \alpha^j J_S \varphi_1^{v^j}(x) \omega(\varphi_1^{v^j}(x), d_x \varphi_1^{v^j} \cdot S) - \bar{\alpha} J_S \varphi_1^{\bar{v}}(x) \omega(\varphi_1^{\bar{v}}(x), d_x \varphi_1^{\bar{v}} \cdot S) d\mu_0 \right| \\
&\leq \int_K |\alpha^j J_S \varphi_1^{v^j}(x)| \left| \omega(\varphi_1^{v_j}(x), d_x \varphi_1^{v_j} \cdot S) - \omega(\varphi_1^{\bar{v}}(x), d_x \varphi_1^{\bar{v}} \cdot S) \right| d\mu_0 \\
&+ \int_K \left| \alpha^j J_S \varphi_1^{v^j}(x) - \bar{\alpha} J_S \varphi_1^{\bar{v}}(x) \right| \left| \omega(\varphi_1^{\bar{v}}(x), d_x \varphi_1^{\bar{v}} \cdot S) \right| d\mu_0
\end{aligned}$$

and using the embedding assumption $W \hookrightarrow C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$

$$\begin{aligned}
&\left|\left(\alpha^j(\varphi_1^{v^j})_\# \mu_0 - \bar{\alpha}(\varphi_1^{\bar{v}})_\# \mu_0 | \omega\right)\right| \\
&\leq \left(\int_K |J_S \varphi_1^{v^j}(x)| d\mu_0 \right) \|\omega\|_{1,\infty} \|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} + C \left[\|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} + |\alpha^j - \bar{\alpha}| \right] \|\omega\|_\infty \\
&\leq C' \left[\|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} + |\alpha^j - \bar{\alpha}| \right] \|\omega\|_W.
\end{aligned}$$

Taking supremum over all $\omega \in W$ with $\|\omega\|_W \leq 1$, we obtain that

$$\|\alpha^j(\varphi_1^{v^j})_\# \mu_0 - \bar{\alpha}(\varphi_1^{\bar{v}})_\# \mu_0\|_{W^*} \leq C' \left[\|(\varphi_1^{v^j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} + |\alpha^j - \bar{\alpha}| \right] \to 0$$

as $j \to \infty$ and therefore $\|\alpha^j(\varphi_1^{v^j})_\# \mu_0 - \mu_{tar}\|_{W^*}^2 \to \|\bar{\alpha}(\varphi_1^{\bar{v}})_\# \mu_0 - \mu_{tar}\|_{W^*}^2$. With the above estimates on the first two terms in the energy (4.1), we conclude that $(\bar{v}, \bar{\alpha})$ is a global minimizer. □

Remark 7. *In fact, one could still prove the existence of minimizers when $\tau = 0$ (by instead assuming $\lambda > 0$) if we make the technical assumption that for any $R > 0$, there exists $\epsilon > 0$ such that for all v with $\|v\|_{L^2([0,1],V)} \leq R$, $\|(\varphi_1^v)_\# \mu_0\|_{W^*} \geq \epsilon$. Indeed, in that case, the boundedness of the varifold distance term:*

$$\begin{aligned}
\|\alpha(\varphi_1^v)_\# \mu_0 - \mu_{tar}\|_{W^*}^2 &= \alpha^2 \|\alpha(\varphi_1^v)_\# \mu_0\|_{W^*}^2 - 2\alpha \langle (\varphi_1^v)_\# \mu_0, \mu_{tar} \rangle_{W^*} + \|\mu_{tar}\|_{W^*}^2 \\
&\geq \alpha^2 \epsilon^2 - 2\alpha\epsilon \|\mu_{tar}\|_{W^*} + \|\mu_{tar}\|_{W^*}^2
\end{aligned}$$

implies that α is bounded as well which allows to use essentially the same argument as in the above proof. Within the class of R-Z kernels (c.f. Definition 8), one can show

that a sufficient condition to recover this technical bound is to take both functions ρ and γ to be strictly positive on their respective domains.

As a matter of fact, one can go a little further by noticing that, as a function of α with v and $\mu(t)$ being fixed, E is quadratic and solving for $\frac{\partial E}{\partial \alpha} = 0$ shows that the optimal α given v can be expressed with respect to the final measure $\mu(1)$ as:

$$\alpha^* = \frac{\frac{\tau}{2} + \lambda \langle \mu(1), \mu_{tar} \rangle_{W^*}}{\frac{\tau}{2} + \lambda \|\mu(1)\|^2_{W^*}}. \tag{4.2}$$

Inserting into (4.1), this now allows to reduce the problem to an optimal control problem with control v:

$$\min_v E(v) \doteq \frac{1}{2} \int_0^1 \|v_t\|^2_V dt + g(\mu(1)), \tag{4.3}$$

where $g(\mu(1)) = \frac{\tau}{2}(\alpha^* - 1)^2 + \lambda \|\alpha^*.\mu(1) - \mu_{tar}\|^2_{W^*}$.

From now on, let v be a minimizer of (4.1) and $q(t)$ the associated optimal trajectory. For numerical consideration, we shall derive some necessary conditions satisfied by such a minimizer for the discrete varifolds case. Following the framw representation discussed in section 2.2.1, we write $\mu_0 = \sum_{i=1}^N r_i \delta_{(x_i, T_i)}$ and, using the representation of the subspaces T_i by frames of d vectors, we can alternatively view the state variable of the optimal control problem as $q = ((x_i, u_i^{(k)})_{1 \leq i \leq N, 1 \leq k \leq d})$, where $\mathrm{Span}(u_i^{(1)}, \dots, u_i^{(d)}) = T_i$ and $|u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}| = \det(u_i^{(k)} \cdot u_i^{(l)}) = r_i$.

The optimality conditions on v can be derived similarly to section 2.2.2, by introducing the Hamiltonian of the problem which is given by:

$$H(p, q, v) \doteq \sum_{i=1}^N \langle p_i^x, v(x_i) \rangle + \sum_{i=1}^N \sum_{k=1}^d \langle p_i^{u_k}, d_{x_i} v(u_i^{(k)}) \rangle - \frac{1}{2} \|v\|^2_V,$$

where $p_i^x, p_i^{u_k} \in \mathbb{R}^n$ denote the costates of the position x and frame vectors $u_i^{(k)}$. By applying the Pontryagin maximum principle, we find that any optimal trajectory

$(x_i(t), u_i^{(k)}(t))$ must satisfy the following Hamiltonian equations:

$$\begin{cases} \dot{x}_i(t) = v_t(x_i(t)) \\ \dot{u}_i^{(k)}(t) = d_{x_i(t)} v_t(u_i^{(k)}(t)) \\ \dot{p}_i^x(t) = -d_{x_i(t)} v_t^T p_i^x(t) - \sum_{k=1}^d d^{(2)}_{x_i(t)} v_t(\cdot, u_i^{(k)}(t))^T p_i^{u_k}(t) \\ \dot{p}_i^{u_k}(t) = -d_{x_i(t)} v_t^T p_i^{u_k}(t) \end{cases} \tag{4.4}$$

and, using the RKHS property of V, the optimal control v is given by:

$$v_t(\cdot) = \sum_{i=1}^N k_V(x_i(t), \cdot) p_i^x(t) + \sum_{k=1}^d \partial_1 k_V(x_i(t), \cdot)(u_i^{(k)}(t), p_i^{u_k}(t)). \tag{4.5}$$

Thus, from the above equations, we obtain that the full energy functional to be minimized can be written as a function of the initial costates, namely

$$\begin{aligned} E(p_i^x(0), p_i^{u_k}(0)) = \frac{1}{2} \sum_{i=1}^N \langle p_i^x(0), v(0)(x_i(0)) \rangle + \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^d \langle p_i^{u_k}(0), d_{x_i(0)} v(0)(u_i^{(k)}(0)) \rangle \\ + \frac{\tau}{2}(\alpha^* - 1)^2 + \lambda \| \alpha^* . \mu(1) - \mu' \|^2_{W^*}, \end{aligned} \tag{4.6}$$

where $\alpha^* . \mu(1) = \sum_{i=1}^N \alpha^* r_i(1) \delta_{(x_i(1), T_i(1))}$ and $T_i(1) = \mathrm{Span}(\{u_i^{(k)}(1)\})$ is obtained from the Hamiltonian equations (4.4) and (4.5).

4.1.2 Numerical implementation

The numerical minimization of the energy (4.6) can be tackled based on an iterative shooting scheme discussed in section 2.2.3. Specifically, given the initial costates $(p_i^x(0), (p_i^{u_k}(0))_{k=1}^d)$ at the current iteration of the algorithm together with the known and fixed initial state variables $(x_i(0), (u_i^{(k)}(0))_{k=1}^d)$, we start by integrating the Hamiltonian equations (4.4) and (4.5) based on an RK4 scheme to obtain the measure $\mu(1)$ at the final time. We then compute α^* with (4.2) from which we obtain the value of the energy (4.6). In order to update the initial costates, we also need the gradient of E which we can directly compute using automatic differentiation. More precisely, the Python implementation leverages the Pytorch library together with the recently developed KeOps library[1]. The latter allows to generate efficient CUDA subroutines

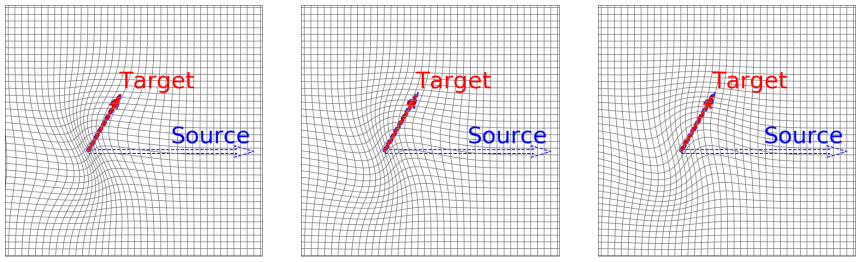
[1] https://www.kernel-operations.io/

for the computation and automatic differentiation of expressions involving positive definite kernels such as the ones appearing in the Hamiltonian equations and in the inner product of W^* given by (1.27). Finally, with E and ∇E being obtained as just explained, the optimization itself is done using the L-BFGS algorithm of the SciPy library.

In what follows, we will refer to this registration algorithm with global density rescaling by the acronym LDDMM+GD. The parameters that need to be set by the user are the kernels k_V and K_W as well as the weighing coefficients λ and τ. The latter controls the relative importance of deformation and mass rescaling in the overall change of density. We illustrate the effect of τ on the simplest example of two single Diracs in Fig. 4-1 (with $n = 2$ and $d = 1$). The optimal diffeomorphism φ_1^v pictured here via the resulting deformed grid shows a combination of a local rotation effect (in order to match the directions of the frame vectors) and of a local compression (to compensate for the difference in mass). The case $\tau = \infty$ corresponds to the pure diffeomorphic registration setting of the previous section. In sharp contrast, when $\tau = 0$, the deformation reduces to only rotating the directional component of the source Dirac while the transformation of mass is entirely done by the rescaling variable α^*. Intermediate values of τ lead to both φ_1^v and α^* contributing to the change in density.

4.2 Local density changes

The model presented in the previous sections is well-suited when a common and global density rescaling effect is expected as the results of Section 4.3 will illustrate but is typically not adapted to the situation of local mass imbalances such as in the case of particular missing parts on the target shape. To tackle this more general case, we develop more generic approaches generalized from the above model, which allow local density change. Instead of the single density rescaling variable α, we consider



(a) $\tau = \infty$ (b) $\tau = 50$ (c) $\tau = 0$

Figure 4-1. Registration between the two Diracs $\mu = r_0\delta_{(x_0,T_0)}$ and $\mu' = r'\delta_{(x',T')}$ with $r_0 = 2$, $r' = 4/5$, $x_0 = x' = (0,0) \in \mathbb{R}^2$ and T, T' are the lines spanned by $u_0 = (1,0)$ and $u' = (\cos(\pi/3), \sin(\pi/3))$. The plotted arrows vectors here represent r_0u_0 and $r'u'$ respectively. The figure illustrates the effect of the choice of τ on the registration, with $\tau = \infty$ corresponding to the pure deformation case (i.e. $\alpha^* = 1$). The optimal density rescaling factors in (b) and (c) are $\alpha^* = 0.6773$ and $\alpha^* = 0.4074$ respectively.

the rescaling function $\alpha(\cdot)$ defined on the ambient space. We shall discuss how to formulate registration models under such idea in detailed in the following subsections.

4.2.1 Simple L^2 model

We first start with a simple extension of the LDDMM+GD model proposed in section 4.1. Instead of assuming constant rescaling factors, we consider rescaling functions living in $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$, the space of positive Borel measurable functions defined on $\mathbb{R}^n$. $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ is a group under pointwise multiplication, and each element α in it can be applied to the varifolds $\mu \in \mathcal{V}_d$ via the action defined by:

$$(\alpha\mu|\omega) = (\mu|\alpha\omega) = \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha(x)\omega(x,T)d\mu(x,T). \tag{4.7}$$

Informally speaking, this action modifies the density of a varifold at each point. Under this setting, the registration of μ_0 to μ_{tar} is formulated as

$$\min_{v,\alpha} E(v,\alpha) \doteq \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \frac{\tau}{2}\int_{\mathbb{R}^n} (\alpha(x)-1)^2 d|\mu_0|(x) + \lambda\|\alpha\mu(1) - \mu_{tar}\|_{W^*}^2, \tag{4.8}$$

subject to $\mu(t) \doteq (\varphi_t^v) \cdot \mu_0$, with the minimization being done over $v \in L^2([0,1], V)$ and $\alpha \in L^2(\mathbb{R}^n, |\mu_0|)$ (the space of L^2 functions on $\mathbb{R}^n$ with respect to the measure

$|\mu_0|$). The penalty on the density rescaling function α is now the distance to the constant function 1 for the L^2 metric weighted by the initial measure μ_0. Once again, the existence of solutions for this model can be shown by the standard approach of calculus of variations:

Proposition 11. *Assume that V is continuously embedded in $C_0^2(\mathbb{R}^n, \mathbb{R}^n)$, W is continuously embedded in $C_0^1(\mathbb{R}^n \times \widetilde{G}_d^n)$ and the kernel of W is C_0-universal, that* $\mathrm{supp}(\mu_0) \subset \mathrm{K}$ *for some compact subset K of $\mathbb{R}^n \times \widetilde{G}_d^n$ and that $\tau > 0$. Then there exist $(v, \alpha) \in L^2([0,1], V) \times L^2(\mathbb{R}^n, |\mu_0|)$ achieving the infimum in (4.8).*

Proof. Let (v^j, η^j) be a minimizing sequence. Since (v^j) is bounded in $L^2([0,1], V)$, we can assume (by extracting a subsequence if necessary) that (v^j) converges weakly to some $\bar{v}$ in $L^2([0,1], V)$ giving that $\|(\varphi_1^{v_j} - \varphi_1^{\bar{v}})|_K\|_{1,\infty} \to 0$ as $j \to +\infty$. Also, since $\tau > 0$, (α^j) is a bounded sequence in $L^2(\mathbb{R}^n, |\mu_0|)$ and thus up to extraction of another subsequence, we may assume that we have weak convergence $\alpha^j \rightharpoonup \bar{\alpha}$ to some $\bar{\alpha} \in L^2(\mathbb{R}^n, |\mu_0|)$.

From the weak lower semicontinuity of the first two terms in the energy, we deduce:

$$\int_0^1 \|\bar{v}\|_V^2 dt \leq \liminf_{j \to +\infty} \int_0^1 \|v^j\|_V^2 dt \tag{4.9}$$

$$\int_{\mathbb{R}^n} (\bar{\alpha}(x) - 1)^2 d|\mu_0|(x) \leq \liminf_{j \to +\infty} \int_{\mathbb{R}^n} (\alpha^j(x) - 1)^2 d|\mu_0|(x) \tag{4.10}$$

In addition, since $\|\varphi_1^{v_j}\|_{1,\infty}$ is bounded and μ_0 is supported in the compact set K, there exists another compact subset $K' \subset \mathbb{R}^n \times \widetilde{G}_d^n$ such that for all $j \in \mathbb{N}$, $\mathrm{supp}(\alpha^j (\varphi_1^{v^j})_\# \mu_0) \subset K'$. Moreover, using the disintegration theorem on the measure μ_0 (c.f. (1.16)), we see that:

$$\begin{aligned} |\alpha^j (\varphi_1^{v^j})_\# \mu_0|(\mathbb{R}^n) &= \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \alpha^j(x) |J_S \varphi_1^{v^j}(x)| d\mu_0(x, S) \\ &= \int_{\mathbb{R}^n} \alpha^j(x) \left(\int_{\widetilde{G}_d^n} |J_S \varphi_1^{v^j}(x)| d\nu_x(S) \right) d|\mu_0|(x) \end{aligned}$$

For the same reason as above, we have that $|J_S \varphi_1^{v^j}(x)|$ is bounded uniformly over $x \in \mathrm{supp}(|\mu_0|)$, $j \in \mathbb{N}$ and $S \in \widetilde{G}_d^n$ from which we get, applying the Cauchy-Schwarz

inequality:

$$|\alpha^j(\varphi_1^{v^j})_\#\mu_0|(\mathbb{R}^n) \leq C\sqrt{\int_{\mathbb{R}^n} \alpha^j(x)^2 d|\mu_0|(x)} = C\|\alpha^j\|_{L^2(\mathbb{R}^n,|\mu_0|)}$$

for some constant $C > 0$. Now since $\|\alpha^j\|_{L^2(\mathbb{R}^n,|\mu_0|)}$ is bounded, we deduce that there exists $M > 0$ such that $|\alpha^j(\varphi_1^{v^j})_\#\mu_0|(\mathbb{R}^n) \leq M$ for all $j \in \mathbb{N}$. In other words, we have obtained that for all j, $\alpha^j(\varphi_1^{v^j})_\#\mu_0$ belong to the space $\mathcal{V}_{d,M,K'}$ defined in Corollary 1. To show the convergence of $\alpha^j(\varphi_1^{v^j})_\#\mu_0$ for the W^* metric, we are thus left to show that it converges for the weak-* topology.

Let $\omega \in C_c(\mathbb{R}^n \times \widetilde{G}_d^n)$.

$$\left(\alpha^j(\varphi_1^{v^j})_\#\mu_0|\omega\right) = \left(\alpha^j(\varphi_1^{\bar{v}})_\#\mu_0|\omega\right) + \left(\alpha^j((\varphi_1^{v^j})_\#\mu_0 - (\varphi_1^{\bar{v}})_\#\mu_0)|\omega\right)$$

Looking at the first term, we have:

$$\begin{aligned}
\left(\alpha^j(\varphi_1^{\bar{v}})_\#\mu_0|\omega\right) &= \int_{\mathbb{R}^n\times\widetilde{G}_d^n} \alpha^j(x) J_S\varphi_1^{\bar{v}}(x)\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)d\mu_0(x,S) \\
&= \int_{\mathbb{R}^n} \alpha^j(x)\left(\int_{\widetilde{G}_d^n} J_S\varphi_1^{\bar{v}}(x)\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)d\nu_x(S)\right)d|\mu_0|(x) \\
&\xrightarrow[j\to\infty]{} \int_{\mathbb{R}^n} \bar{\alpha}(x)\left(\int_{\widetilde{G}_d^n} J_S\varphi_1^{\bar{v}}(x)\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)d\nu_x(S)\right)d|\mu_0|(x) \\
&= \left(\bar{\alpha}(\varphi_1^{\bar{v}})_\#\mu_0|\omega\right)
\end{aligned}$$

where the convergence in the third row follows from the weak convergence of α^j to $\bar{\alpha}$ in $L^2(\mathbb{R}^n, |\mu_0|)$ and the fact the function between parentheses is measurable and bounded on $\mathbb{R}^n$ and thus in $L^2(\mathbb{R}^n, |\mu_0|)$. As for the second term, we can expand it and see that:

$$\begin{aligned}
&\left|\left(\alpha^j((\varphi_1^{v^j})_\#\mu_0 - (\varphi_1^{\bar{v}})_\#\mu_0)|\omega\right)\right| \\
&\leq \int_K |\alpha^j(x)|\left|J_S\varphi_1^{v^j}(x)\omega(\varphi_1^{v^j}(x), d_x\varphi_1^{v^j}\cdot S) - J_S\varphi_1^{\bar{v}}(x)\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)\right| d\mu_0(x,S) \\
&\leq \|\alpha^j\|_{L^2(\mathbb{R}^n,|\mu_0|)}\cdot \\
&\quad \left(\int_K \left|J_S\varphi_1^{v^j}(x)\omega(\varphi_1^{v^j}(x), d_x\varphi_1^{v^j}\cdot S) - J_S\varphi_1^{\bar{v}}(x)\omega(\varphi_1^{\bar{v}}(x), d_x\varphi_1^{\bar{v}}\cdot S)\right|^2 d\mu_0(x,S)\right)^{1/2}
\end{aligned}$$

Now, $\|\alpha^j\|_{L^2(\mathbb{R}^n,|\mu_0|)}$ is bounded and by convergence of $\varphi_1^{v^j}$ to $\varphi_1^{\bar{v}}$ in $\|\cdot\|_{1,\infty}$ and uniform continuity of ω, we deduce that $\left(\alpha^j((\varphi_1^{v^j})_\#\mu_0 - (\varphi_1^{\bar{v}})_\#\mu_0)|\omega\right) \to 0$. Therefore, we have shown that $\alpha^j(\varphi_1^{v^j})_\#\mu_0 \overset{*}{\rightharpoonup} \bar{\alpha}(\varphi_1^{\bar{v}})_\#\mu_0$. By Corollary 1, this implies that $\alpha^j(\varphi_1^{v^j})_\#\mu_0 \xrightarrow{\|\cdot\|_{W^*}} \bar{\alpha}(\varphi_1^{\bar{v}})_\#\mu_0$. As a result,

$$\lim_{j\to\infty} \|\alpha^j(\varphi_1^{v^j})_\#\mu_0 - \mu_{tar}\|_{W^*}^2 = \|\bar{\alpha}(\varphi_1^{\bar{v}})_\#\mu_0 - \mu_{tar}\|_{W^*}^2 \tag{4.11}$$

Combining (4.9), (4.10) and (4.11), we obtain:

$$E(\bar{v},\bar{\alpha}) \leq \lim\inf_{j\to\infty} E(v^j,\alpha^j)$$

and thus $(\bar{v},\bar{\alpha})$ is a minimizer.

□

Considering now the discrete varifold case in view of the numerical implementation, let $\mu_0 = \sum_{i=1}^N r_i\delta_{(x_i,T_i)}$ and $\mu_{tar} = \sum_{j=1}^M w_j\delta_{(y_j,S_j)}$. In this discrete setting, we may consider the rescaling functions to be simple functions of the form $\alpha(x) = \sum_{i=1}^N \alpha_i \mathbf{1}_{x_i}(x)$, which support coincide with the support of μ_0 and the measure mass scaling becomes $\alpha\mu_0 = \sum_{i=1}^N \alpha_i r_i \delta_{(x_i,T_i)}$. The discrete registration problem is written as follows:

$$\min_{v,\alpha\in\mathbb{R}_+^N} E(v,\alpha) \doteq \frac{1}{2}\int_0^1 \|v(t)\|_V^2 dt + \frac{\tau}{2}\sum_{i=1}^N r_i(\alpha_i - 1)^2 + \lambda\|\alpha\cdot\mu(1) - \mu'\|_{W^*}^2. \tag{4.12}$$

Problem (4.12) can be solved in similar fashion as in Section 4.1.2. The Hamiltonian equations (4.4) and (4.5) still hold. However, the optimality equations on the α_i's are not as straightforward to exploit as (4.2). Indeed, differentiating E with respect to α_k and obtain the following linear system:

$$\begin{aligned} 0 = \frac{\partial E}{\partial \alpha_k} &= 2\lambda\sum_{i=1}^N K(x_k,T_k,x_i,T_i)r_k(1)r_i(1)\alpha_i - 2\lambda\sum_{j=1}^M K(x_k,T_k,y_j,S_j)r_k(1)w_j \\ &\quad + \delta r_k(0)\alpha_k - \delta r_k(0). \end{aligned}$$

The above linear system can be written in the matrix form,

$$(2\lambda\boldsymbol{A} + \tau\mathbf{Diag}(\boldsymbol{r}(0)))\boldsymbol{\alpha} = 2\lambda\boldsymbol{b} + \tau\boldsymbol{r}(0),$$

where

$$\boldsymbol{A} \doteq (K(x_i, T_i, x_j, T_j) r_i(1) r_j(1))_{i,j},\ \boldsymbol{b} \doteq (\sum_{j=1}^{M} K(x_i, T_i, y_j, S_j) r_i(1) w_j)_i$$
$$\boldsymbol{\alpha} \doteq (\alpha_1, \cdots, \alpha_N),\ \boldsymbol{r}(t) \doteq (r_1(t), \cdots, r_N(t)),$$

So the solution is

$$\boldsymbol{\alpha} = (2\lambda \boldsymbol{A} + \tau \mathbf{Diag}(\boldsymbol{r}(0)))^{-1} (2\lambda \boldsymbol{b} + \tau \boldsymbol{r}(0)) .$$

In practice, we instead jointly optimize over the initial costates $(p_i^x(0), (p_i^{u_k}(0))_{k=1}^d)$ together with $\boldsymbol{\alpha} = (\alpha_i)$ using L-BFGS, with the gradient of the energy with respect to each α_i being computed by automatic differentiation. We will denote this diffeomorphic registration under local density changes by the acronym LDDMM+L2.

4.2.2 LDDMM-Fisher-Rao model

In the classical shape analysis setting, shape spaces are modelled as homogeneous spaces under the action of a group of transformations. Metrics between shapes can then be defined from the right invariant metrics on the transformation groups. However, in the previous sections, our approaches first let the source varifold μ_0 be transported by a pure geometric transformation φ_1^v, then apply the rescaling factor α to the deformed varifold at end time. Such approaches make it difficult to define shape spaces and metrics due to its lack of symmetry. The goal of this section is to propose an alternative model more in line with the Grenander's classical shape theory, that defines the weight rescaling through a dynamical process and introduces shape metrics from group actions on varifolds.

As in section (4.2.1), we consider rescaling functions $\alpha \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ that transforms a varifold μ according to (4.7). We can combine the applications of the diffeomorphism group $\mathrm{Diff}(\mathbb{R}^n)$ and of $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ on varifolds by semi-direct product. We define a homomorphism $\Psi : \mathrm{Diff}(\mathbb{R}^n) \mapsto \mathrm{Aut}(\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+))$, which sends elements φ to automorphisms $\Psi_\varphi(\alpha) \doteq \alpha \circ \varphi$ of $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$. The group $\mathrm{Diff}(\mathbb{R}^n) \ltimes_\Psi \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ of semi-direct

product between $\mathrm{Diff}(\mathbb{R}^n)$ and $\mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ can be defined by the following group law:

$$(\varphi_1, \alpha_1) \cdot (\varphi_2, \alpha_2) \doteq (\varphi_1 \circ \varphi_2, (\alpha_1 \circ \varphi_2)\alpha_2),$$

and it is straightforward to verify that the identity is $(\mathrm{id}, 1)$ and the inverse of the element (φ, α) is $(\varphi^{-1}, \frac{1}{\alpha \circ \varphi^{-1}})$. A natural left action of the group $\mathrm{Diff}(\mathbb{R}^n) \ltimes_\Psi \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ on the space of varifolds $\mathcal{V}_d$ can be then defined as follows:

$$(\varphi, \alpha) \cdot \mu \doteq \varphi_\#(\alpha\mu) = (\alpha \circ \varphi^{-1})\varphi_\#\mu, \tag{4.13}$$

which corresponds to first rescaling μ by α then deforming $\alpha\mu$ by φ, or equivalently first deforming μ by φ then recaling the deformed varifold by α with the change of coordinate φ^{-1}.

Let's consider $\mu_0 \in \mathcal{V}_d$ and velocities $(v, \eta) \in L^2([0,1], V \times L^2(\mathbb{R}^n, \mathbb{R}))$. We can generate smooth paths $(\varphi_t^v, \alpha_t^\eta)$ in $\mathrm{Diff}(\mathbb{R}^n) \ltimes_\Psi \mathcal{B}(\mathbb{R}^n, \mathbb{R}_+)$ from the following ODE:

$$\begin{cases} \partial_t \varphi_t = v_t \circ \varphi_t \\ \partial_t \alpha_t = \alpha_t \eta_t, \\ \varphi_0 = \mathrm{id},\ \alpha_0 = 1. \end{cases}, \tag{4.14}$$

for $t \in [0,1]$. From that path, we get a path $\mu_t \doteq (\varphi_t^v, \alpha_t^\eta) \cdot \mu_0$ in $\mathcal{V}_d$ defined by the group action (4.13). We define the cost for such path joining μ_0 and μ_1 as follows:

$$C_{\mu_0}(v, \eta) = \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \frac{\tau}{2}\int_0^1 \int_{\mathbb{R}^n} \eta_t^2 \circ (\varphi_t^v)^{-1}(x) d|\mu_t|(x) dt. \tag{4.15}$$

The first term in the above functional measures the kinetic energy from the vector field v and the second term measures the energy of growth rate η. The spatial integral inside the second term can be written as

$$\begin{aligned} \int_{\mathbb{R}^n} \eta_t^2 \circ (\varphi_t^v)^{-1}(x) d|\mu_t|(x) &= \int_{\mathbb{R}^n \times \tilde{G}_d^n} \eta_t^2 \circ (\varphi_t^v)^{-1}(x) d\mu_t(x, T) \\ &= \int_{\mathbb{R}^n \times \tilde{G}_d^n} \eta_t^2(x) J_T \varphi_t^v(x) \alpha_t(x) d\mu_0(x, T). \end{aligned}$$

When $v_t = 0, \forall t \in [0,1]$, $\mu_t = \alpha_t \mu_0$, the spatial integral in the second term is the Fisher-Rao metric tensor, and minimizing the second term gives the Fisher-Rao metric[70, 98–100] between measures $|\mu_0|$ and $|\mu_1|$.

Let Θ be a homogeneous space generated by a varifold $\mu_0 \in \mathcal{V}_d$ with compact support *i.e.* $\Theta \doteq \{(\varphi, \alpha) \cdot \mu_0 | (\varphi, \alpha) \in G_V \ltimes_\Psi \mathcal{B}(\mathbb{R}^n, \mathbb{R}^+)\}$. Now for μ and $\mu' \in \Theta$, we define

$$d_\Theta(\mu, \mu') \doteq (\inf_{v,\eta}\{C_\mu(v, \eta) \mid (\varphi_1^v, \alpha_1^\eta) \cdot \mu = \mu'\})^{1/2}, \tag{4.16}$$

which is a true distance as stated in the following Theorem:

Theorem 14. *The function d_Θ given by (4.16) defines a distance on Θ, i.e. it is symmetric, satisfies the triangle inequality and $d_\Theta(\mu, \mu') = 0$ if and only if $\mu = \mu'$.*

Proof. The symmetry can be proved by the time reversal argument. Let $\mu, \mu' \in \Theta$ and $(v, \eta) \in L^2([0, 1], V \times L^2(\mathbb{R}^n, |\mu_0|))$ such that $\mu' = (\alpha_1^\eta \circ (\varphi_1^v)^{-1})(\varphi_1^v)_\# \mu$ and we denote by $\mu_t \doteq (\varphi_t^v, \alpha_t^\eta) \cdot \mu_0$. If we define

$$(\bar{v}_t, \bar{\eta}_t) \doteq -(v_{1-t}, \eta_{1-t} \circ (\varphi_1^v)^{-1}),$$

then $\varphi_t^{\bar{v}} \circ \varphi_1^v = \varphi_{1-t}^v$ and

$$\begin{aligned}
&(\alpha_t^{\bar{\eta}} \circ (\varphi_t^{\bar{v}})^{-1})(\alpha_1^\eta \circ (\varphi_{1-t}^v)^{-1}) \\
&= \left(e^{-\int_0^t \eta_{1-s}(\cdot)ds} \circ (\varphi_1^v)^{-1} \circ (\varphi_t^{\bar{v}})^{-1}\right)\left(e^{\int_0^1 \eta_s(\cdot)ds} \circ (\varphi_{1-t}^v)^{-1}\right) \\
&= e^{\int_0^{1-t} \eta_s(\cdot)ds} \circ (\varphi_{1-t}^v)^{-1} \\
&= \alpha_{1-t}^\eta \circ (\varphi_{1-t}^v)^{-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\alpha_t^{\bar{\eta}} \circ (\varphi_t^{\bar{v}})^{-1}(\varphi_t^{\bar{v}})_\# \mu' &= (\alpha_t^{\bar{\eta}} \circ (\varphi_t^{\bar{v}})^{-1})(\alpha_1^\eta \circ (\varphi_{1-t}^v)^{-1})(\varphi_t^{\bar{v}} \circ \varphi_1^v)_\# \mu \\
&= \alpha_{1-t}^\eta \circ (\varphi_{1-t}^v)^{-1}(\varphi_{1-t}^v)_\# \mu \\
&= \mu_{1-t}.
\end{aligned}$$

The energy for this reverse trajectory is given by:

$$C_{\mu'}(\bar{v}, \bar{\eta}) = \frac{1}{2}\int_0^1 \|v_{1-t}\|_V^2 dt + \frac{\tau}{2}\int_0^1 \int_{\mathbb{R}^n} \eta_{1-t}^2 \circ (\varphi_{1-t}^v)^{-1} d|\mu_{1-t}|(x)dt.$$

This gives the symmetry.

For triangular inequality, let's consider μ, μ' and μ'' in Θ such that $d_\Theta(\mu,\mu')$, $d_\Theta(\mu',\mu'')>0$. Assuming $(v,\eta),(v',\eta')\in L^2([0,1],V\times L^2(\mathbb{R}^n,|\mu_0|))$ with

$$\mu_1=\mu' \text{ and } \mu_1'=\mu'',$$

where

$$\mu_t\doteq(\varphi_t^v)_\#(\alpha_t^\eta\mu) \text{ and } \mu_t'\doteq(\varphi_t^{v'})_\#(\alpha_t^{\eta'}\mu').$$

Let $a,b>1$ such that $1/a+1/b=1$ and denote $u_t\doteq\min\{at,1\}$ and $u_t'\doteq\max\{b(t-1/a),0\}$. Defining

$$(\bar v_t,\bar\eta_t)\doteq a(v_{u_t},\eta_{u_t})\mathbf{1}_{[0,1/a)}(t)+b(v'_{u_t'},\eta'_{u_t'}\circ\varphi_1^v)\mathbf{1}_{[1/a,1]}(t).$$

One can easily check that

$$\begin{aligned}&\varphi_t^{\bar v}=\varphi_{u_t'}^{v'}\circ\varphi_{u_t}^v,\ \alpha_t^{\bar\eta}=(\alpha_{u_t'}^{\eta'}\circ\varphi_1^v)\alpha_{u_t}^\eta\\&\text{and }\bar\mu_t\doteq(\varphi_t^{\bar v})_\#(\alpha_t^{\bar\eta}\mu)=\mu_{u_t}\mathbf{1}_{[0,1/a)}(t)+\mu_{u_t'}\mathbf{1}_{[1/a,1]}(t).\end{aligned}$$

From above, we can further obtain

$$\begin{aligned}C_\mu(\bar v,\bar\eta)&=a\left(\frac12\int_0^{1/a}\left(\|v_{u_t}\|_V^2+\frac\tau2\int_{\mathbb{R}^n}\eta_{u_t}^2\circ(\varphi_{u_t}^v)^{-1}(x)d|\mu_{u_t}|(x)\frac{du_t}{dt}\right)dt\right)\\&\quad+b\left(\frac12\int_{1/a}^1\left(\|v'_{u_t'}\|_V^2+\frac\tau2\int_{\mathbb{R}^n}(\eta'_{u_t'})^2\circ(\varphi_{u_t'}^{v'})^{-1}(x)d|\mu'_{u_t'}|(x)\frac{du_t'}{dt}\right)dt\right)\\&=aC_\mu(v,\eta)+bC_{\mu'}(v',\eta').\end{aligned}$$

Taking $a^*=1+\frac{C_{\mu'}(v',\eta')^{1/2}}{C_\mu(v,\eta)^{1/2}}$, we can check that

$$C_\mu(\bar v,\bar\eta)^{1/2}=C_\mu(v,\eta)^{1/2}+C_{\mu'}(v',\eta')^{1/2},$$

and this impies the triangular inequality.

If $d_{\mathcal{M}}(\mu,\mu')=0$, then there exists a minimizing sequence (v^j,η^j) such that $\mu'=(\varphi_1^{v^j})_\#(\alpha_1^{\eta^j}\mu)$, $\forall j$. Note that $\sup_{t\in[0,1]}\|\varphi_t^{v^j}-Id\|_{1,\infty}\to0$ on any compact subset

of $\mathbb{R}^n$ from the fact that $\int_0^1 \|v_t^j\|_V^2 dt \to 0$ (7.16) in [21]. Let A denote the support of μ, it's easy to see that for large n,

$$\sup_{t\in[0,1]} \sup_{(x,T)\in A} |J_T \varphi_t^{v^j}(x) - 1| < \frac{1}{2}.$$

From the inequality above and the sequence (v^j, η^j) is minimizing, we have

$$\int_0^1 \int_{\mathbb{R}^n} (\eta_t^j(x))^2 \alpha_t^{\eta^j}(x) d|\mu|(x) \leq C \int_0^1 \int_{\mathbb{R}^n \times \tilde{G}_d^n} (\eta_t^j(x))^2 \alpha_t^{\eta^j}(x) J_T \varphi_t^{v^j}(x) d\mu(x,T) \to 0,$$

as $j \to \infty$. This implies that $\alpha_1^{\eta^j}|\mu|$ converges to $|\mu|$ in the Fisher-Rao metric, and

$$\int_{\mathbb{R}^n} (\sqrt{\alpha_1^{\eta^j}} - 1)^2 d|\mu|(x) \to 0$$

from the L^2 isometry of the Fisher-Rao metric [70]. This also implies that $\int_{\mathbb{R}^n} |\alpha_1^{\eta_j}| d\mu \to |\mu|(\mathbb{R}^n) < \infty$ and $\alpha_1^{\eta_j} \to 1$ $|\mu|$-*a.e.* up to a subsequence. From Theorem 4.6.2 in [101], $\alpha_1^{\eta_j}$ converges to 1 in $L^1(|\mu|)$. For any bounded continuous function ω,

$$\begin{aligned}(\mu - \mu'|\omega) = &\underbrace{\int_{\mathbb{R}^n \times \tilde{G}_d^n} \omega(x,T) d\mu(x,T) - \int_{\mathbb{R}^n \times \tilde{G}_d^n} \omega(\varphi_1^{v^j}(x), d_x \varphi_1^{v^j} \cdot T) J_T \varphi_1^{v^j}(x) d\mu(x,T)}_{I} \\ &+ \underbrace{\int_{\mathbb{R}^n \times \tilde{G}_d^n} \omega(\varphi_1^{v^j}(x), d_x \varphi_1^{v^j} \cdot T)(1 - \alpha_1^{\eta^j}(x)) J_T \varphi_1^{v^j}(x) d\mu(x,T)}_{II}.\end{aligned}$$

It's clear that $I \to 0$ from bounded convergence theorem. Also, from the fact that $\alpha_1^{\eta^j}$ converges to 1 in $L^2(|\mu|)$ and the uniform control of Jacobian above, we obtain that

$$|II| \leq C \int_{\mathbb{R}^n} |1 - \alpha^{\eta^j}(x)| d|\mu|(x) \to 0.$$

This implies that $\mu = \mu'$. □

With the regularization term defined in (4.15), we can formulate the inexact registration of μ_0 to μ_{tar} as the following optimization problem:

$$\min_{v,\eta} E(v,\eta) \doteq C_{\mu_0}(v,\eta) + \lambda \|\mu_1 - \mu_{tar}\|_{W^*}^2,$$

where $\mu_t \doteq (\varphi_t^v, \alpha_t^\eta) \cdot \mu_0$ and $(\varphi_t^v, \alpha_t^\eta)$ is obtained from (4.14). As in sections 4.1 and 4.2.1, we shall derive optimality equations for the discrete case. Assuming $\mu_0 = \sum_{i=1}^N r_i \delta_{(x_i,T_i))}$

and from the frame representation, we can write the state variable of the optimal control problem as $q_0 = ((x_i, u_i^{(k)})_{1\le i\le N, 1\le k\le d})$, where $\mathrm{Span}(u_i^{(1)}, \dots, u_i^{(d)}) = T_i$ and $|u_i^{(1)} \wedge \cdots \wedge u_i^{(d)}| = \det(u_i^{(k)} \cdot u_i^{(l)}) = r_i$. Using this representation, the dynamical equation $\mu_t = (\varphi_t^v, \alpha_t^\eta) \cdot \mu_0$ is equivalent to the following action on the frame model:

$$\begin{aligned} q_t &\doteq (\varphi_t^v, \alpha_t^\eta) \cdot q_0 \\ &= (x_i(t), u_i^{(1)}(t), \cdots, u_i^{(d)}(t))_{i=1}^N \\ &= (\varphi_t^v(x_i(0)), \alpha_{i,t}^{1/d} d_{x_i(0)}\varphi_t^v(u_i^{(1)}(0)), \cdots, \alpha_{i,t}^{1/d} d_{x_i(0)}\varphi_t^v(u_i^{(d)}(0)))_{i=1}^N, \end{aligned}$$

where $\alpha_{i,t} \doteq \alpha_t^\eta(x_i(t)) \in \mathbb{R}_+$. From above, we can derive the state equations:

$$\begin{cases} \dot{x}_i(t) = v_t(x_i(t)) \\ \dot{u}_i^{(k)}(t) = d_{x_i(t)} v_t(u_i^{(k)}(t)) + \frac{1}{d}\eta_{i,t} u_i^{(k)}(t) \end{cases}$$

where $\eta_{i,t} \doteq \eta_t(x_i(t)) \in \mathbb{R}$. The energy functional for this discrete optimal control problem can be written as:

$$E(v, \eta) \doteq \frac{1}{2}\int_0^1 \|v_t\|_V^2 dt + \frac{\delta}{2}\sum_{i=1}^N \int_0^1 \eta_{i,t}^2 r_i(t) dt + \lambda \|\mu^{q_1} - \mu^{tar}\|_{W^*}^2, \tag{4.17}$$

where $r_i(t) \doteq |u_i^{(1)}(t) \wedge \cdots \wedge u_i^{(d)}(t)|$ and μ^q denote the varifold corresponding to the state q. The Hamiltonian for this problem are given by:

$$H(p, q, v, \eta) \doteq \sum_{i=1}^N \langle p_i^x, v(x_i)\rangle + \sum_{i=1}^N \sum_{k=1}^d \langle p_i^{u_k}, \left(d_{x_i} v + \frac{\eta_i}{d} I_d\right) u_i^{(k)}\rangle - \frac{1}{2}\|v\|_V^2 - \frac{\tau}{2}\sum_{i=1}^N \eta_i^2 r_i.$$

From the Pontryagin maximum principle, any optimal trajectories $(x_i(t), u_i^{(k)}(t))$ must satisfy the following Hamiltonian equations:

$$\begin{cases} \dot{x}_i(t) = v_t(x_i(t)) \\ \dot{u}_i^{(k)}(t) = d_{x_i(t)} v_t(u_i^{(k)}(t)) + \frac{\eta_{i,t}}{d} u_i^{(k)(t)} \\ \dot{p}_i^x(t) = -d_{x_i(t)} v_t^T p_i^x(t) - \sum_{k=1}^d d_{x_i(t)}^{(2)} v_t(\cdot, u_i^{(k)}(t))^T p_i^{u_k}(t) \\ \dot{p}_i^{u_k}(t) = -(d_{x_i(t)} v_t^T + \frac{\eta_{i,t}}{d} I_d) p_i^{u_k}(t) + \frac{\tau}{2}\frac{\eta_{i,t}^2}{r_i(t)} \sum_{\ell=1}^d C_i^{\ell k}(t) u_i^{(\ell)}(t) \end{cases}$$

where $\boldsymbol{C}_i(t) = (C_i^{\ell k}(t))_{1\le \ell, k\le d}$ is the cofactor matrix of the gramian $(\langle u_i^\ell(t), u_i^k(t)\rangle)_{\ell,k}$.

The optimal controls (v_t, η_t) are given by

$$\begin{cases} v_t(\cdot) = \sum_{i=1}^N k_V(x_i(t), \cdot) p_i^x(t) + \sum_{k=1}^d \partial_1 k_V(x_i(t), \cdot)(u_i^{(k)}(t), p_i^{u_k}(t)) \\ \eta_{i,t} = \frac{1}{\tau d r_i} \sum_{k=1}^d \langle p_i^{u_k}(t), u_i^{(k)}(t)\rangle. \end{cases} \tag{4.18}$$

And the reduced Hamiltonian has the following form:

$$H(p,q) \doteq \frac{1}{2}\left\{\sum_{i=1}^{N}\langle p_i^x, v(x_i)\rangle + \sum_{i=1}^{N}\sum_{k=1}^{d}\langle p_i^{u_k}, d_{x_i}v(u_i^{(k)})\rangle + \tau\sum_{i=1}^{N}\eta_i^2 r_i\right\}$$
$$= \frac{1}{2}\left\{\sum_{i=1}^{N}\langle p_i^x, v(x_i)\rangle + \sum_{i=1}^{N}\sum_{k=1}^{d}\langle p_i^{u_k}, d_{x_i}v(u_i^{(k)})\rangle + \frac{1}{\tau d^2}\sum_{i=1}^{N}\frac{(\sum_{k=1}^{d}\langle p_i^{u_k}, u_i^{(k)}\rangle)^2}{r_i}\right\}$$

where v and η are given by (4.18).

Our python implementation of minimizing (4.17) is performed with the iterative shooting scheme as described in section 4.1.2, so we don't repeat details here. In what follows, we will denote this registration framework by the acronym LDDMM+FR.

4.3 Numerical results

As proof-of-concept, we present a few results of the above LDDMM+GD, LDDMM+L2 and LDDMM+FR algorithms applied to 2D and 3D shapes ($n = 2$ or 3), specifically discrete curves ($d = 1$) or surfaces ($d = 2$). Those shapes are converted to varifolds as explained in Section 1.3.1. For the purpose of visualization however, we shall plot the shapes rather than their associated measures in $\mathcal{V}_d$ and display the source shape's time evolution along the estimated deformation path φ_t^v. In all experiments, the deformation kernel k_V is a Gaussian and the kernel K_W defining the metric on $\mathcal{V}_d$ is chosen among the class of kernels discussed in [55], specifically as the tensor product of a Gaussian kernel on $\mathbb{R}^n$ and the Binet kernel between d-dimensional subspaces, i.e. for $T = \mathrm{Span}(u^{(k)})$ and $T' = \mathrm{Span}(u'^{(k)})$ the positive kernel given by $\det((u^{(k)} \cdot u'^{(l)})_{k,l})^2$. We also set τ to a small value so as to put only minimal constraints on the estimation of α.

Fiber bundles. For our first set of simulations, we consider white matter fiber tracts taken from the publicly available ISMRM 2015 Tractography Challenge repository [2]. In Fig. 4-2, we show the result of registering a single template curve onto the

[2] http://www.tractometer.org/ismrm_2015_challenge/

posterior commissure (CP) bundle containing 365 distinct curves. As can be seen on the first row, regular LDDMM registration generates a folding of the source curve in an attempt to compensate for the difference in total mass. The LDDMM+GD algorithm on the other hand leads to a deformed curve that matches the average geometry of the bundle with an estimated $\alpha^* = 348.80$ consistent with the density of curves in the target. Note that α^* is in fact smaller than 365 which accounts for the spatial spreading and fanning of the bundle. We also consider the registration between two different fiber tracts: the anterior commissure (CA) and fornix which are made of respectively 431 and 2151 individual curves. Once again, standard diffeomorphic registration (not shown here for the sake of space) induces important artifactual folding effects in contrast with the LDDMM+GD registration result of Fig. 4-3.

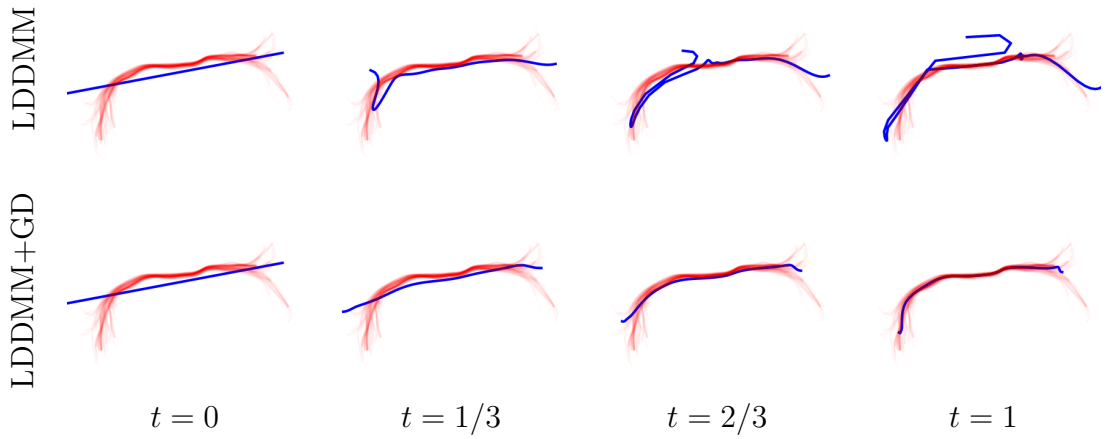


Figure 4-2. Registration of single curve to CP fiber bundle (365 curves). The second row shows the deformation at intermediate times for the proposed LDDMM+GD algorithm where the estimated density rescaling is $\alpha^* = 348.80$.

Sparse shapes. Another feature of the LDDMM+GD approach is its robustness under sparse and incomplete observation of the target surface. This is illustrated on the example of hippocampi surfaces (data provided with the Keops library) in Fig. 4-4 where sparse targets are synthetically generated by keeping only a small number of random faces from the full ground truth target mesh. This mass imbalance

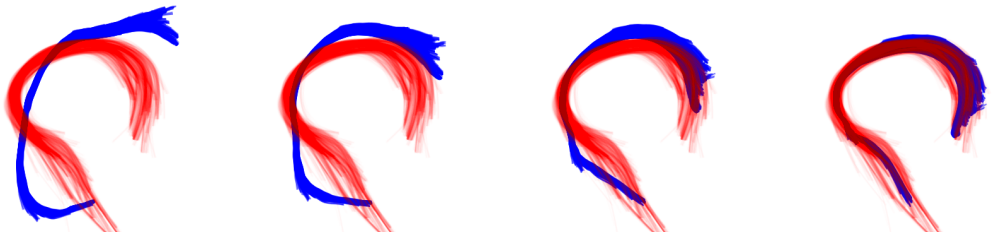


Figure 4-3. Registration of CA fiber bundle (431 curves) to Fornix fiber bundle (2151 curves). The estimated α^* is 4.18, close to the fiber density ratio of the bundles.

results in severe shrinking and twisting of the registered surface estimated with standard LDDMM while LDDMM+GD recovers a surface close to the ground truth and automatically estimates (through α^*) the sparsity rate with good accuracy.

Curves with densities. The LDDMM+GD model is not well-suited in the situation where submanifolds with non-constant densities. In figure 4-5, we consider a synthetic example of matching a circle with constant density 1 to an ellipse with piecewise constant density function with values .5, .75, 1.25 and 1.75. Such a case requires the registration process to increase and decrease density function values on different components of the deformed curve, which cannot be done by standard LDDMM and LDDMM+GD models, while the LDDMM+L2 and LDDMM+FR models allow such local mass changes and are able to estimate the density values on different components. As shown on the first row, the colors represent the densities on the target, and estimated from LDDMM+L2 and LDDMM+FR respectively. Note that the dynamics of the estimated values in the LDDMM+L2 model were computed by the linear interpolation $(1-t) + t\alpha_i$. In the second row, we plot α_i and $\alpha_{i,1}$ as functions of the angle and the histograms of these values. We observed that the function $\alpha_{i,1}$ from LDDMM+FR model oscillates more than α_i from the LDDMM+L2 model.

Partial matching. The other situation that the LDDMM+GD is however not well-suited is that an isolated part of the target shape is locally missing, as with the

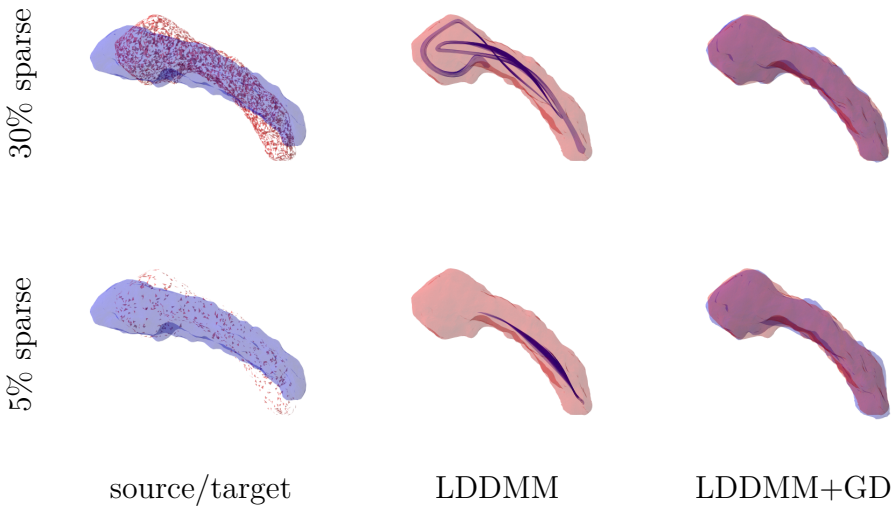


Figure 4-4. Registration between two hippocampus surfaces: source (in blue) and target (in red) which has been randomly subsampled to 30% (first row) and 5% (second row) of its total number of triangles. The resulting deformed surface at $t = 1$ obtained with LDDMM (second column) and LDDMM+GD (third column) is compared to the fully sampled target surface. For LDDMM+GD, we obtain $\alpha^* = 0.29$ and $\alpha^* = 0.049$. The Hausdorff distance to the ground truth is 1.3755 (30%) and 1.7720 (5%) for the LDDMM+GD model versus 5.0485 and 16.1139 for LDDMM.

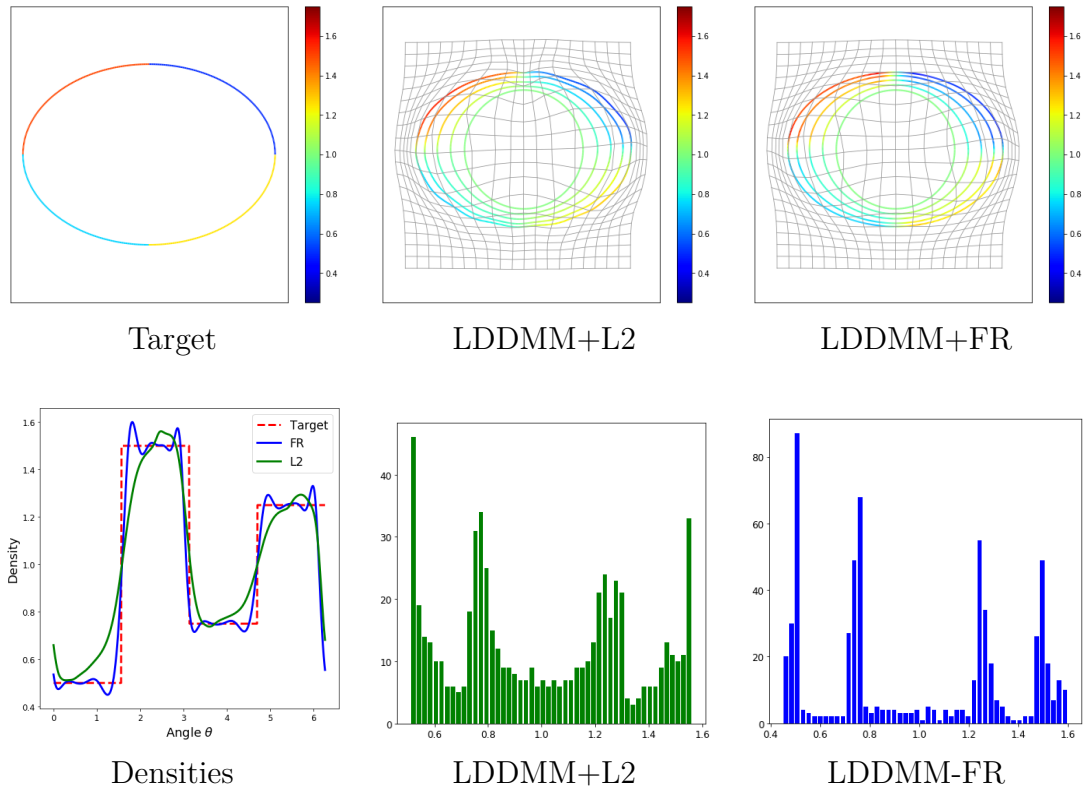


Figure 4-5. Registration of a circle with uniform density 1 to an ellipse with piecewise constant density, .5, .75, 1.25 and 1.75. The first column shows the target ellipse with colors corresponding the density at each location and the plot of the densities as functions of angle for the target and the mass rescaling results obtained from LDDMM+L2 and LDDMM+FR models. The second row presents the evolution of the deformed curve and the histogram of the mass rescaling from the LDDMM+L2 model. In the top one plot, the dynamics of mass rescaling is computed from linear intepolation between 1 and α_i and the grids visualize the diffeomorphism at $t = 1$. The third column shows the deformed curve, α_{i_t} and the histogram of $\alpha_{i,1}$ of LDDMM+FR model.

simulated example of Fig. 4-6 in which we artificially remove a subregion of the hippocampus surface. In such a case, the LDDMM+L2 and LDDMM+FR algorithms allow for local mass changes and is able to estimate, alongside the deformation, the corresponding missing region on the source shape as shown on the third and fourth rows, where the colors represent the values of the α_i's and $\alpha_{i,t}$'s at the different locations. In contrast, such missing regions can adversely affect registration under the standard LDDMM model. This is evidenced quantitatively by the closer proximity

(measured for the usual Hausdorff distance) of the matched surface to the ground truth (i.e. complete) target for our proposed LDDMM+L2 and LDDMM+FR approaches. The dynamics of α_i in the LDDMM+L2 is obtained by linear interpolation between 1 and α_i. We also observed that $\alpha_{i,1}$ in LDDMM+FR oscillate more than α_i in LDDMM+L2, which was also seen in Fig. 4-5. From the third row in Fig. 4-6, we can see that these oscillations come from the fact that the LDDMM+FR tends to shrink part of triangular meshes and put higher values of $\alpha_{i,1}$ on these meshes.

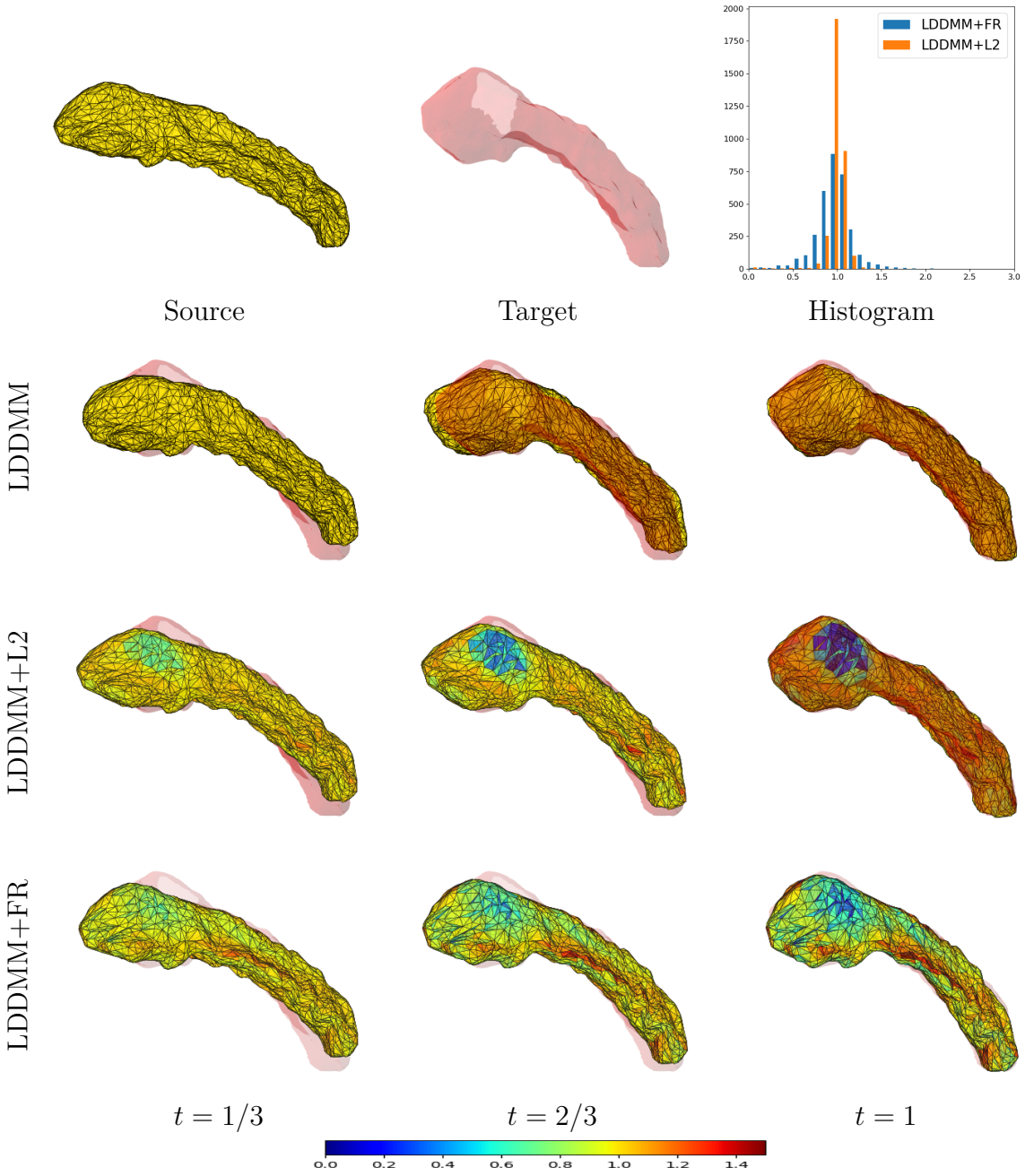


Figure 4-6. Registration of two hippocampi surfaces with a missing subregion obtained with the LDDMM+L2, LDDMM+FR, and standard LDDMM approach. The colors on the second and third rows correspond to the values of the mass rescaling α_i and $\alpha_{i,t}$ at each location. One can notice small differences in terms of overlap between the final shape and the ground truth target around that subregion. The Hausdorff distance to the ground truth is 1.5714 (LDDMM+L2), 1.6291 (LDDMM+FR) and 1.9789 (LDDMM).

Chapter 5

Conclusions and perspectives

In this thesis, we proposed a registration framework between varifolds that goes beyond the previous restrictions of such models to the registration of discrete or smooth submanifolds of $\mathbb{R}^n$. To achieve so, we studied a general class of distances between oriented varifolds based on reproducing kernels and derived a deformation model on the space $\mathcal{V}_d$, which are combined into an optimal control formulation of the registration problem between any two varifolds. We also examined the possibility to couple this approach with a quantization/compression methodology in order to eventually tackle the registration problem, in practice, on discrete varifolds with a relatively low number of Dirac masses. In the last part of this thesis, we extend our varifolds registration model allowing global/local density changes in order to tackle comparison and matching of imbalanced shapes.

We showed that first of all this setting leads to an equivalent yet alternative formulation to the diffeomorphic registration of rectifiable sets such as continuous or discrete curves and surfaces; the resulting higher-order Hamiltonian systems in our model provides richer local patterns for the deformations but at the price of a higher numerical cost. From an application standpoint, however, the main advantage we expect from this framework is that it applies very naturally to more general geometric objects, in particular to typical situations where well-defined and reliable meshes are not available. We gave a taste of it through some of the examples of Section 2.3,

although future work on a larger scale will be needed in order to evaluate such benefits more thoroughly.

Besides the cases mentioned here, there are also several types of data that could constitute interesting test applications for this setting. This includes for instance high-angular resolution diffusion MRI in which the data is effectively modeled as spatially distributed orientation probability distribution functions consistent with the Young measure representation of varifolds in (1.16), or the case of contrast-invariant image registration c.f. [77]. In Chapter 4, we show that our extended density changes models can handle challenging registrations including fiber bundles containing significant differences in numbers of curves and hippocampus surfaces with missing parts. We also observed that the LDDMM+FR model somehow tends to shrink parts of the shapes while increasing the densities on those parts. This could be improved in the future by imposing an additional penalty term to enforce the local alignment of masses.

At the theoretical level, there are several questions left open by this work which we believe can constitute interesting tracks for future work. One is to study the possibility of extending all or part of the results of Section 3.2 to more general kernel metrics (in particular currents) and determining tighter quantization error bounds. Moreover, the registration model at play in this paper is based on the pushforward group action of $\mathrm{Diff}(\mathbb{R}^n)$ on $\mathcal{V}_d$. Yet, other group actions could be have been considered, as briefly evoked in Section 1.3.2, that involve different choices of reweighing factor, for which we could expect very different properties of the solutions to the registration problem.

Some additional work on the numerical side is likely needed for potential future applications to large scale databases, most notably to generalize this work to the estimation of means and atlases over populations of many high resolution shapes. Indeed, as we pointed out, even with the ability to compress the size of varifolds in the registration pipeline using the quantization approach, the higher complexity of the dynamical equations involved in the registration model has a non-negligible numerical

toll. This could be improved in the future by using more efficient computational schemes for the repeated evaluations of sums of kernels and derivative of kernels appearing in the Hamiltonian equations, possibly along the lines of fast multiple methods. Building deep learning models could be one potential approach to address this problem. Deep neural networks refer to a family of parametrized functions that use multiple layers to extract features from the input data or to approximate general functions. While having achieved great success in machine learning, deep learning has gradually drawn attention in the community of shape analysis. In the recent work [102], the authors proposed a preliminary image matching scheme which generate the registration map through convolution neural network (CNN). However, there is no guarantee that the obtained mapping is a diffeomorphism and it does not provide a dynamic path of transformation as LDDMM. Among existing deep learning architectures, residual neural networks (also called ResNets) are the most relevant ones to deformation analysis since the mapping blocks in ResNets can be viewed as Euler's discretization schemes for ODEs. In the work [103], the authors have tried to make use of ResNets to generate diffeomorphic flows with some mild loss to enforce the inverse consistency of the resulting maps. And [104] proposed a ResNet based diffeomorphic registration framework 3D shapes, and 3D point clouds. These works showed the applicability of neural networks for generating deformations and registrations. By choosing proper architectures and loss functions, it hints at the possibility of designing a more computationally efficient varifold registration framework based on neural networks.

Other than directly using neural networks to generate deformations, there is an alternative way to apply deep learning to reduce the computation time. Recall that with the geodesic shooting scheme, the geodesics of diffeomorphisms in G_V are encoded by initial momenta or vector fields. The works [15, 105], take advantage of such encoding to train neural networks to predict optimal initial momenta/vector

fields for image registrations. Once such networks are trained, the registration can be done immediately by solving the forward geodesic equations with the predicted initial momenta. We explored this approach for curve matching in the specific context of speech processing [106, 107] as part of a collaborative work with the Neural Systems Analysis Laboratory (NSA Lab) at Johns Hopkins University. In these works, the objective was to tackle the problem of emotion voice conversion, namely, given a neutral utterance, to train a model to transform the emotional content from one to the other while preserving the semantic or speaker information. The strategy implemented in those works consisted in deforming the pitch contour of a source emotional utterance to that of a target emotional utterance using LDDMM curve registration, then reconstruct the voice from the deformed pitch contour. As described previously, registrations of curves were done by predicting the initial momenta of the shooting formulation with a trained neural network. To simplify the implementation of the neural network, pitch contours were reparametrized so that they all contain the same number of points. Although this is rather a simple situation for curve matching, similar ideas could be explored in the future to provide more generic frameworks for varifold matching.

Another possible track of interest is to develop a range of efficient and robust algorithms for conversion between the representations of discrete varifolds, point clouds, and discrete surfaces (See figure 5-1). In figure 2-5, we showed that our varifold registration framework can be applied to match noisy point clouds. To apply our algorithm, we first need to represent the point sets as discrete varifolds, which requires estimating local tangent spaces and local masses. As stated in section 2.3, this is done by the GMRA algorithm [84], which provides local partitions with estimated tangent planes. However, this algorithm does not estimate local masses for each partition. To resolve this, we project points in each partition to its corresponding tangent plane and obtain the mass by roughly estimating the area of the region occupied by

projected points on the plane. These procedures are pretty cumbersome and require a large number of points to obtain decent performance. In order to make the varifold representation more applicable to point clouds analysis, a next step could be to design models and algorithms that can conveniently estimate varifold representations from the point clouds. Recently, [108] proposed to estimate masses from point clouds by regularizing measures via convolutions. Another possibility could be to even to consider a continuous model $x \mapsto (r(x), T(x))$, which takes the point positions as input and outputs the estimated weights and tangent spaces at those positions and then develop a prediction framework using state of the art machine learning algorithms. Besides the problem of varifold estimation, another related interesting problem could be surface reconstruction from discrete varifolds. There are well-established methods for simplification or compression of shapes represented as general measures such as currents or varifolds and algorithms to obtain triangulated surfaces from (noisy) point clouds [109–111]. We have also discussed how to represent discrete curves and surfaces as discrete varifolds in section 1.3.1. However, there is no straightforward way to reconstruct curves or surfaces from discrete currents/varifolds representations. Robust algorithms remain to be developed that would allow to recover a curve or surface closest to a given discrete varifolds in terms of kernel metrics.

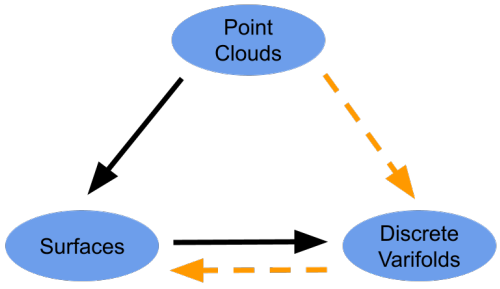


Figure 5-1. Diagram of algorithms to change from one shape representation to the other. While there are already methods to construct surface from point clouds and represent surface as discrete varifolds, robust algorithms for estimating discrete varifolds from point clouds and recovering surfaces from discrete varifolds are still to be developed.

Lastly, although this thesis focused on the LDDMM model for deformations, we want to mention that it is not the only shape registration framework that varifold representation can be paired with. In our recent work [112], we proposed a new surface matching algorithm that, in a spirit similar to Chapter 3, combines specific elastic metrics and varifold distances. Let us briefly summarize the approach. In elastic shape analysis, the shape space of surfaces is defined as the quotient parametrized surfaces by the reparametrized group. Consider the space of parametrized surfaces as the set $\mathcal{I}$ of all oriented immersions q from a compact manifold Ω (with or without boundary) into $\mathbb{R}^3$, then the shape space is defined as $M \doteq \mathcal{I}/\mathrm{Diff}(\Omega)$ which consists of the unparametrized surfaces, that is, the equivalence classes $[q] = \{q \circ \varphi | \varphi \in \mathrm{Diff}(\Omega)\}$. The square root normal field (SRNF) of an oriented parametrized surface $q \in \mathcal{I}$ [113, 114] is the unit normal field multiplied by the square root of the Riemannian area, which can be written as

$$N_q \doteq \frac{q_x \wedge q_y}{\sqrt{|q_x \wedge q_y|}},$$

in local coordinates. The SRNF distance between two parametrized surfaces q_0 and q_1 is defined as the L^2 distance between the associated SRNFs:

$$\mathrm{dist}_N(q_0, q_1) \doteq \left(\int_\Omega |N_{q_0} - N_{q_1}|^2 \right)^{1/2}.$$

It is easy to see that the above is $\mathrm{Diff}(\Omega)$-invariant, i.e.,

$$\mathrm{dist}_N(q_0 \circ \varphi, q_1 \circ \varphi) = \mathrm{dist}_N(q_0, q_1). \tag{5.1}$$

It then a classical result that one can recover a (pseudo-)distance on the shape space M by quotienting out the reparametrizations, namely:

$$\mathrm{dist}_N([q_0], [q_1]) \doteq \inf_{\varphi \in \mathrm{Diff}(\Omega)} \mathrm{dist}_N(q_0, q_1 \circ \varphi). \tag{5.2}$$

Despite dist_N being very simple to compute, the main difficulty in estimating (5.2) is that it requires finding an optimal reparametrization $\varphi \in \mathrm{Diff}(\Omega)$.

A possible approach to address this issue is to rely on the fact that varifold kernel distances d_{W^*} are insensitive to surface parametrization:

$$d_{W^*}(q_0, q_1) = d_{W^*}(q_0 \circ \varphi_0, q_1 \circ \varphi_1). \tag{5.3}$$

for any $\varphi_0, \varphi_1 \in \mathrm{Diff}(\Omega)$. With (5.3), problem (5.2) is equivalent to the following constrained optimization problem:

$$\min_{\tilde{q}_1 \in \mathcal{I}} \mathrm{dist}_N(q_0, \tilde{q}_1), \text{ subject to } d_{W^*}(\tilde{q}_1, q_1) = 0.$$

Relaxation of the constraint using a Lagrange multiplier λ yields the following variational problem:

$$\min_{\tilde{q}_1 \in \mathcal{I}} \mathrm{dist}_N(q_0, \tilde{q}_1) + \lambda d_{W^*}(\tilde{q}_1, q_1).$$

The above formulation is a computationally efficient registration framework since, on the one hand, it provides a simple registration functional to be optimized, on the other hand, it allows us to bypass the difficulty of finding optimal parametrizations.

References

1. Grenander, U. & Miller, M. I. Computational anatomy: An emerging discipline. *Quarterly of applied mathematics* **56,** 617–694 (1998).
2. Ma, J., Miller, M. I., Trouvé, A. & Younes, L. Bayesian template estimation in computational anatomy. *NeuroImage* **42,** 252–261 (2008).
3. Durrleman, S., Pennec, X., Trouvé, A. & Ayache, N. *A forward model to build unbiased atlases from curves and surfaces* in *2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy* (2008), 68–79.
4. Durrleman, S., Fillard, P., Pennec, X., Trouvé, A. & Ayache, N. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage* **55,** 1073–1090 (2011).
5. Lombaert, H. *et al.* Human atlas of the cardiac fiber architecture: study on a healthy population. *IEEE transactions on medical imaging* **31,** 1436–1447 (2012).
6. Oishi, K. *et al.* Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participants. *Neuroimage* **46,** 486–499 (2009).
7. Feng, J., Tang, X., Tang, M., Priebe, C. & Miller, M. *Metric space structures for computational anatomy* in *International Workshop on Machine Learning in Medical Imaging* (2013), 123–130.
8. Tang, X. *et al.* Shape abnormalities of subcortical and ventricular structures in mild cognitive impairment and Alzheimer's disease: detecting, quantifying, and predicting. *Human brain mapping* **35,** 3701–3725 (2014).
9. Durrleman, S. *et al.* Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *NeuroImage* **101,** 35–49 (2014).
10. Niethammer, M., Huang, Y. & Vialard, F.-X. *Geodesic regression for image time-series* in *International conference on medical image computing and computer-assisted intervention* (2011), 655–662.
11. Fishbaugh, J., Prastawa, M., Gerig, G. & Durrleman, S. *Geodesic shape regression in the framework of currents* in *International Conference on Information Processing in Medical Imaging* (2013), 718–729.
12. Fishbaugh, J., Durrleman, S., Prastawa, M. & Gerig, G. Geodesic shape regression with multiple geometries and sparse parameters. *Medical image analysis* **39,** 1–17 (2017).

13. Tang, X. *et al.* Regional subcortical shape analysis in premanifest Huntington's disease. *Human brain mapping* **40,** 1419–1433 (2019).

14. Mumford, D. Pattern Theory: the Mathematics of Perception. *Proceedings of the International Congress of Mathematicians* **III,** 1–21 (2002).

15. Yang, X., Kwitt, R., Styner, M. & Niethammer, M. Quicksilver: Fast predictive image registration–a deep learning approach. *NeuroImage* **158,** 378–396 (2017).

16. Weber, G. W. *et al.* Virtual anthropology: the digital evolution in anthropological sciences. *Journal of physiological anthropology and Applied Human Science* **20,** 69–80 (2001).

17. Zhang, Y. & Bajaj, C. Finite element meshing for cardiac analysis. *Univ. of Texas at Austin: ICES Technical Report* (2004).

18. Thompson, D. W. *On Growth and Form* (Cambridge university press, 1917).

19. Grenander, U. *General pattern theory-A mathematical study of regular structures* (Clarendon Press, 1993).

20. Grenander, U. & Miller, M. I. *Pattern theory: from representation to inference* (Oxford university press, 2007).

21. Younes, L. *Shapes and diffeomorphisms* 2nd ed. (Springer, 2019).

22. Lee, J. M. *Riemannian manifolds: an introduction to curvature* (Springer Science & Business Media, 2006).

23. Carmo, M. P. d. *Riemannian geometry* (Birkhäuser, 1992).

24. Kriegl, A. & Michor, P. W. *The convenient setting of global analysis* (American Mathematical Soc., 1997).

25. Michor, P. W. & Mumford, D. A zoo of diffeomorphism groups on $\mathbb{R}^n$. *Annals of Global Analysis and Geometry* **44,** 529–540 (2013).

26. Holm, D. D. *Geometric Mechanics* (World Scientific Publishing Company, 2008).

27. Beg, M. F., Miller, M. I., Trouvé, A. & Younes, L. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision* **61** (2005).

28. Trouvé, A. Diffeomorphisms groups and pattern matching in image analysis. *International journal of computer vision* **28,** 213–221 (1998).

29. Dupuis, P., Grenander, U. & Miller, M. I. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of applied mathematics,* 587–600 (1998).

30. Arguillère, S., Trélat, E., Trouvé, A. & Younes, L. Shape deformation analysis from the optimal control viewpoint. *Journal de Mathématiques Pures et Appliquées* **104,** 139–178 (2015).

31. Miller, M. & Qiu, A. The emerging discipline of Computational Functional Anatomy. *NeuroImage* **45,** 16–39 (2009).

32. Miller, M., Trouvé, A. & Younes, L. Hamiltonian Systems and Optimal Control in Computational Anatomy: 100 Years Since D'Arcy Thompson. *Annu Rev Biomed Eng* **7,** 447–509 (2015).

33. Aronszajn, N. Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68,** 337–404 (1950).

34. Berlinet, A. & Thomas-Agnan, C. *Reproducing kernel Hilbert spaces in probability and statistics* (Springer Science & Business Media, 2011).

35. Hastie, T., Tibshirani, R. & Friedman, J. *The elements of statistical learning* (Springer, 2001).

36. Micheli, M. & Glaunès, J. Matrix-valued kernels for shape deformation analysis. *Geometry, Imaging and Computing* **1,** 57–139 (2014).

37. Zhou, D.-X. Derivative reproducing properties for kernel methods in learning theory. *Journal of computational and Applied Mathematics* **220,** 456–463 (2008).

38. Pontryagin, L., Boltyanskii, V., Gamkrelidze, R. & Mishchenko, E. *The Mathematical Theory of Optimal Processes* (John Wiley & Sons, 1962).

39. Arguillère, S. The general setting for shape analysis. *arXiv preprint arXiv:1504.01767* (2015).

40. Arguillere, S. Sub-Riemannian geometry and geodesics in Banach manifolds. *The Journal of Geometric Analysis* **30,** 2897–2938 (2020).

41. Arguillere, S. & Trélat, E. Sub-Riemannian structures on groups of diffeomorphisms. *Journal of the Institute of Mathematics of Jussieu* **16,** 745–785 (2017).

42. Vialard, F.-X., Risser, L., Rueckert, D. & Cotter, C. Diffeomorphic 3D Image Registration via Geodesic Shooting Using an Efficient Adjoint Calculation. *International Journal of Computer Vision* **97,** 229–241 (2012).

43. Miller, M. I., Trouvé, A. & Younes, L. Geodesic shooting for computational anatomy. *Journal of mathematical imaging and vision* **24,** 209–228 (2006).

44. Miller, M. I., Younes, L. & Trouvé, A. Diffeomorphometry and geodesic positioning systems for human anatomy. *Technology* **2,** 36–43 (2014).

45. Charon, N. & Trouvé, A. The varifold representation of nonoriented shapes for diffeomorphic registration. *SIAM Journal on Imaging Sciences* **6,** 2547–2580 (2013).

46. Joshi, S. C. & Miller, M. I. Landmark matching via large deformation diffeomorphisms. *IEEE transactions on image processing* **9,** 1357–1370 (2000).

47. Glaunès, J., Trouvé, A. & Younes, L. Diffeomorphic Matching of Distributions: A New Approach for Unlabelled Point-Sets and Sub-Manifolds Matching. *Computer Vision and Pattern Recognition (CVPR)* **2,** 712–718 (2004).

48. Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B. & Smola, A. A kernel two-sample test. *The Journal of Machine Learning Research* **13,** 723–773 (2012).

49. Sriperumbudur, B. K., Fukumizu, K. & Lanckriet, G. R. Universality, characteristic kernels and RKHS embedding of measures. *Journal of Machine Learning Research* **12,** 2389–2410 (2011).

50. Simon, L. *Lecture notes on geometric measure theory* (Australian national university, 1983).

51. Evans, L. *Measure theory and fine properties of functions* (Routledge, 2018).

52. Vaillant, M. & Glaunes, J. *Surface matching via currents* in *Biennial International Conference on Information Processing in Medical Imaging* (2005), 381–392.

53. Glaunès, J., Qiu, A., Miller, M. & Younes, L. Large Deformation Diffeomorphic Metric Curve Mapping. *Int J Comput Vis* **80,** 317–336 (2008).

54. Charon, N. *Analysis of geometric and functional shapes with extensions of currents: applications to registration and atlas estimation* PhD thesis (École normale supérieure de Cachan-ENS Cachan, 2013).

55. Kaltenmark, I., Charlier, B. & Charon, N. *A general framework for curve and surface comparison and registration with oriented varifolds* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), 3346–3355.

56. Bauer, M., Bruveris, M., Charon, N. & Møller-Andersen, J. A relaxed approach for curve matching with elastic metrics. *ESAIM: Control, Optimisation and Calculus of Variations* **25,** 72 (2019).

57. Roussillon, P. & Glaunès, J. A. Kernel metrics on normal cycles and application to curve matching. *SIAM Journal on Imaging Sciences* **9,** 1991–2038 (2016).

58. Roussillon, P. & Glaunès, J. A. *Surface matching using normal cycles* in *International Conference on Geometric Science of Information* (2017), 73–80.

59. Roussillon, P. & Glaunès, J. A. Representation of surfaces with normal cycles and application to surface registration. *Journal of Mathematical Imaging and Vision* **61,** 1069–1095 (2019).

60. Charon, N., Charlier, B., Glaunès, J., Gori, P. & Roussillon, P. in *Riemannian geometric statistics in medical image analysis* 441–477 (Elsevier, 2020).

61. Young, L. C. Generalized surfaces in the calculus of variations. *Annals of mathematics* **43,** 84–103 (1942).

62. Almgren, F. J. *Plateau's problem: an invitation to varifold geometry* (American Mathematical Soc., 1966).

63. Allard, W. K. On the first variation of a varifold. *Annals of mathematics,* 417–491 (1972).

64. Morgan, F. *Geometric measure theory: a beginner's guide* (Elsevier Science, 2008).

65. Lee, J. M. *Introduction to smooth manifolds* (Springer Science & Business Media, 2013).

66. Ambrosio, L., Fusco, N. & Pallara, D. *Functions of bounded variation and free discontinuity problems* (Oxford : Clarendon Press, 2000).

67. Federer, H. *Geometric measure theory* (Springer, 2014).

68. Villani, C. *Optimal transport: old and new* (Springer Science & Business Media, 2008).

69. Piccoli, B. & Rossi, F. Generalized Wasserstein Distance and its Application to Transport Equations with Source. *Archive for Rational Mechanics and Analysis* **211,** 335–358 (2014).

70. Chizat, L., Peyré, G., Schmitzer, B. & Vialard, F.-X. An Interpolating Distance Between Optimal Transport and Fisher-Rao Metrics. *Foundations of Computational Mathematics* **18,** 1–44 (2018).

71. Bogachev, V. I. *Measure theory* (Springer Science & Business Media, 2007).

72. Feydy, J., Charlier, B., Vialard, F.-X. & Peyré, G. *Optimal Transport for Diffeomorphic Registration* in *Medical Image Computing and Computer Assisted Intervention* (2017), 291–299.

73. Peyré, G., Cuturi, M., *et al.* Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning* **11,** 355–607 (2019).

74. Hamm, J. & Lee, D. D. *Grassmann discriminant analysis: a unifying view on subspace-based learning* in *Proceedings of the 25th international conference on Machine learning* (2008), 376–383.

75. Wolf, L. & Shashua, A. Learning over sets using kernel principal angles. *Journal of Machine Learning Research* **4,** 913–931 (2003).

76. Harandi, M. T., Salzmann, M., Jayasumana, S., Hartley, R. & Li, H. *Expanding the family of grassmannian kernels: An embedding perspective* in *European Conference on Computer Vision* (2014), 408–423.

77. Hsieh, H.-W. & Charon, N. *Diffeomorphic registration of discrete geometric distributions* in *Mathematics of Shapes and Applications* (World Scientific, 2020), 45–74.

78. Hsieh, H.-W. & Charon, N. Metrics, quantization and registration in varifold spaces. *Foundations of Computational Mathematics,* 1–45 (2021).

79. Sommer, S., Nielsen, M., Darkner, S. & Pennec, X. Higher-Order Momentum Distributions and Locally Affine LDDMM Registration. *SIAM Journal on Imaging Sciences* **6,** 341–367 (2013).

80. Charlier, B., Feydy, J. & Glaunès, J. *Keops (Software)* https://www.kernel-operations.io/. 2018.

81. Overton, M. *HANSO: hybrid algorithm for non-smooth optimization 2.2* https://cs.nyu.edu/overton/software/hanso/. 2016.

82. Miller, M. *et al.* Network Neurodegeneration in Alzheimer's Disease via MRI Based Shape Diffeomorphometry and High-Field Atlasing. *Frontiers in Bioengineering and Biotechnology* **3,** 54 (2015).

83. Ardekani, S. *et al.* Matching Sparse Sets of Cardiac Image Cross-Sections Using Large Deformation Diffeomorphic Metric Mapping Algorithm. *Statistical Atlases and Computational Models of the Heart* (2011).

84. Allard, W. K., Chen, G. & Maggioni, M. Multi-scale geometric methods for data sets II: Geometric multi-resolution analysis. *Applied and Computational Harmonic Analysis* **32,** 435–462 (2012).

85. Carmeli, C., De Vito, E., Toigo, A. & Umanita, V. Vector valued reproducing kernel Hilbert spaces and universality. *Analysis and Applications* **8,** 19–61 (2010).

86. Berg, C., Christensen, J. P. R. & Ressel, P. *Harmonic analysis on semigroups: theory of positive definite and related functions* (Springer, 1984).

87. Graf, S. & Luschgy, H. *Foundations of quantization for probability distributions* (Springer, 2007).

88. Kloeckner, B. Approximation by finitely supported measures. *ESAIM: Control, Optimisation and Calculus of Variations* **18,** 343–359 (2012).

89. Buet, B., Leonardi, G. P. & Masnou, S. Discretization and approximation of surfaces using varifolds. *Geometric Flows* **3,** 28–56 (2018).

90. Durrleman, S., Pennec, X., Trouvé, A. & Ayache, N. Statistical models of sets of curves and surfaces based on currents. *Medical image analysis* **13,** 793–808 (2009).

91. Durrleman, S. *Statistical models of currents for measuring the variability of anatomical curves, surfaces and their evolution* PhD thesis (Université Nice Sophia Antipolis, 2010).

92. Gori, P. *et al.* Parsimonious Approximation of Streamline Trajectories in White Matter Fiber Bundles. *IEEE Transactions on Medical Imaging* **PP** (2016).

93. Chauffert, N., Ciuciu, P., Kahn, J. & Weiss, P. A Projection Method on Measures Sets. *Constructive Approximation* **45,** 83–111 (2017).

94. Charlier, B., Nardi, G. & Trouvé, A. The matching problem between functional shapes via a BV-penalty term: a Γ-convergence result. *arXiv preprint arXiv:1503.07685* (2015).

95. Arguillère, S., Miller, M. & Younes, L. Diffeomorphic Surface Registration with Atrophy Constraints. *SIAM Journal on Imaging Sciences* **9,** 975–1003 (2016).

96. Richardson, C. & Younes, L. Computing metamorphoses between discrete measures. *Journal of Geometric Mechanics* **5,** 131 (2013).

97. Hsieh, H.-W. & Charon, N. *Diffeomorphic Registration with Density Changes for the Analysis of Imbalanced Shapes* in *International Conference on Information Processing in Medical Imaging* (2021), 31–42.

98. Rao, C. R. Information and the accuracy attainable in the estimation of statistical parameters. *Reson. J. Sci. Educ* **20,** 78–90 (1945).

99. Ay, N., Jost, J., Vân Lê, H. & Schwachhöfer, L. Information geometry and sufficient statistics. *Probability Theory and Related Fields* **162,** 327–364 (2015).

100. Bauer, M., Bruveris, M. & Michor, P. W. Uniqueness of the Fisher–Rao metric on the space of smooth densities. *Bulletin of the London Mathematical Society* **48,** 499–506 (2016).

101. Durrett, R. *Probability: theory and examples* (Cambridge university press, 2019).

102. Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J. & Dalca, A. V. *An unsupervised learning model for deformable medical image registration* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), 9252–9260.

103. Rousseau, F., Drumetz, L. & Fablet, R. Residual networks as flows of diffeomorphisms. *Journal of Mathematical Imaging and Vision* **62,** 365–375 (2020).

104. Amor, B. B., Arguillère, S. & Shao, L. ResNet-LDDMM: Advancing the LDDMM Framework Using Deep Residual Networks. *arXiv preprint arXiv:2102.07951* (2021).

105. Yang, X., Kwitt, R., Styner, M. & Niethammer, M. *Fast predictive multimodal image registration* in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (2017), 858–862.

106. Shankar, R., Hsieh, H.-W., Charon, N. & Venkataraman, A. Automated emotion morphing in speech based on diffeomorphic curve registration and highway networks. *Proc. Interspeech 2019* (2019).

107. Shankar, R., Hsieh, H.-W., Charon, N. & Venkataraman, A. Multi-speaker Emotion Conversion via Latent Variable Regularization and a Chained Encoder-Decoder-Predictor Network. *arXiv preprint arXiv:2007.12937* (2020).

108. Buet, B. & Rumpf, M. Mean curvature motion of point cloud varifolds. *arXiv preprint arXiv:2010.09419* (2020).

109. Bærentzen, J. A., Gravesen, J., Anton, F. & Aanæs, H. *Guide to computational geometry processing: foundations, algorithms, and methods* (Springer Science & Business Media, 2012).

110. Choi, G. P.-T., Ho, K. T. & Lui, L. M. Spherical conformal parameterization of genus-0 point clouds for meshing. *SIAM Journal on Imaging Sciences* **9,** 1582–1618 (2016).

111. Sharp, N. & Ovsjanikov, M. *Pointtrinet: Learned triangulation of 3d point sets* in *European Conference on Computer Vision* (2020), 762–778.

112. Bauer, M., Charon, N., Harms, P. & Hsieh, H.-W. A numerical framework for elastic surface matching, comparison, and interpolation. *International Journal of Computer Vision,* 1–20 (2021).

113. Jermyn, I. H., Kurtek, S., Klassen, E. & Srivastava, A. *Elastic shape matching of parameterized surfaces using square root normal fields* in *European conference on computer vision* (2012), 804–817.

114. Kurtek, S., Klassen, E., Gore, J. C., Ding, Z. & Srivastava, A. Elastic geodesic paths in shape space of parameterized surfaces. *IEEE transactions on pattern analysis and machine intelligence* **34,** 1717–1730 (2011).

Appendix

Proofs:

Proof of Proposition 2

From the area formula (Theorem 8), we can see that

$$\begin{aligned}(\phi_{\#}\mu_X|\omega) &= \int_X \omega(\phi(x), T_{\phi(X)}(\phi(x)))J_{T_xX}\phi(x)d\mathcal{H}^d(x)\\ &= \int_{\phi(X)} \omega(y, T_{\phi(X)}(y)d\mathcal{H}^d(y))\\ &= (\mu_{\phi(X)}|\omega),\end{aligned}$$

for all $\omega \in C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$. □

Proof of Proposition 3

Since d_{BL} metrizes the narrow topology, it suffices to show that μ_i converges to μ in the narrow topology. Let ω be a bounded continuous function defined on $\mathbb{R}^n \times \widetilde{G}_d^n$ and $\varepsilon > 0$. By the tightness property, we may choose a compact set $K \subset \mathbb{R}^n \times \widetilde{G}_d^n$ such that $\mu(K^c) + \sup_i \mu_i(K^c) < \varepsilon/2\|\omega\|_\infty$. Let B be an open ball that contains K. Define

$$\eta(x,T) \doteq \begin{cases} \omega(x,T), & \text{if } (x,T) \in K \\ 0, & \text{if } (x,T) \in B^c \end{cases}$$

From Tietz extension theorem, there exists a continuous extension $\tilde{\omega}$ of η on $\mathbb{R}^n \times \widetilde{G}_d^n$

such that $\tilde{\omega}|_K = \omega|_K$ and $\tilde{\omega} \in C_c(\mathbb{R}^n \times \widetilde{G}_d^n)$. This implies that

$$\begin{aligned}\left|\int_{\mathbb{R}^n\times\widetilde{G}_d^n} \omega d(\mu_i - \mu)\right| &\leq \int_{\mathbb{R}^n\times\widetilde{G}_d^n} |\omega - \tilde{\omega}| d(\mu_i + \mu) + \left|\int_{\mathbb{R}^n\times\widetilde{G}_d^n} \tilde{\omega} d(\mu_i - \mu)\right| \\ &\leq \left|\int_{\mathbb{R}^n\times\widetilde{G}_d^n} \tilde{\omega} d(\mu_i - \mu)\right| + \varepsilon.\end{aligned}$$

Taking lim sup on both sides, we see that

$$\limsup_{i\to\infty} \left|\int_{\mathbb{R}^n\times\widetilde{G}_d^n} \omega d(\mu_i - \mu)\right| < \varepsilon.$$

Since ε is arbitrary, we obtain that μ_i converges to μ in the narrow topology. □

Proof of Theorem 4

Since k^{pos} and k^G are bounded, k is also bounded. Therefore, we have the following inequality:

$$\begin{aligned}\|\omega\|_\infty &= \sup_{x,T} |(\delta_{(x,T)}|\omega)| \leq \|\delta_{(x,T)}\|_{W^*}\|\omega\|_W \\ &= \sqrt{k((x,T),(x,T))}\|\omega\|_W \leq \sqrt{\|k\|_\infty}\|\omega\|_W\end{aligned}$$

Next, we show that $W \subset C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$. Let W_0 denote the set of all finite linear combinations of functions in the form $k((x,T),\cdot)$. Recall that W_0 is a dense subspace of W. Let $\omega \in W$, there exists a sequence $\omega_n \in W_0$ such that $\lim_{m\to\infty} \|\omega - \omega_m\|_W = 0$. Therefore we can obtain

$$\|\omega - \omega_m\|_\infty \leq \sqrt{\|k\|_\infty}\|\omega - \omega_m\|_W \to 0$$

as $m \to \infty$. Since $(C_0(\mathbb{R}^n \times \widetilde{G}_d^n), \|\cdot\|_\infty)$ is a complete metric space, $\omega \in C_0(\mathbb{R}^n \times \widetilde{G}_d^n)$.

□

Proof of Proposition 5

Recall that for all $\phi \in \mathrm{Diff}(\mathbb{R}^n)$, $g(\phi) = \lambda\|\phi_\#\mu_0 - \mu_{tar}\|^2_{W^*}$ which we may rewrite as

$$g(\phi) = \lambda(\phi_\#\mu_0|K_W(\phi_\#\mu_0 - 2\mu_{tar})) + \lambda\|\mu_{tar}\|^2_{W^*}.$$

Thus, the variation with respect to ϕ in the Banach space $\mathcal{B}$ writes

$$\partial_\phi g(\phi) = \partial_\phi(\phi_\# \mu_0 | \omega_0)$$

where $\omega_0 \doteq 2\lambda K_W(\phi_\#\mu_0 - \mu_{tar}) \in W$. Moreover

$$(\phi_\#\mu_0|\omega_0) = \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \omega_0(\phi(x), d_x\phi \cdot T) J_T\phi(x) d\mu_0(x,T).$$

Taking the variation with respect to ϕ along any $u \in C_0^1(\mathbb{R}^n, \mathbb{R}^n)$, we obtain:

$$\begin{aligned}(\partial_\phi g(\phi)|u) = &\int_{\mathbb{R}^n \times \widetilde{G}_d^n} \partial_x \omega_0(\phi(x), d_x\phi \cdot T) \cdot u(x) J_T\phi(x) d\mu_0(x,T) \\ &+ \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \partial_T \omega_0(\phi(x), d_x\phi \cdot T) \cdot (d_x u|_T) J_T\phi(x) d\mu_0(x,T) \\ &+ \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \omega_0(\phi(x), d_x\phi \cdot T) . \mathrm{div}_T u(x) . J_T\phi(x) d\mu_0(x,T) \end{aligned} \tag{.1}$$

where the last term follows from the differentiation of Gram determinant matrices while the notation ∂_T in the second term is a shortcut notation for differentiation on the Grassmannian which we do not explicit further here, we however refer to the similar computations done in [45] and to the developments in Section 2.2 for more details. For the first term, we can rely on the Young measure decomposition $\mu_0 = |\mu_0| \otimes \nu_x$ introduced in (1.16) which gives:

$$\begin{aligned}(1) &= \int_{\mathbb{R}^n} \tilde{\alpha}(\phi, x) \cdot u(x) \; d|\mu_0|(x), \\ &\text{where} \;\; \tilde{\alpha}(\phi, x) = \int_{\widetilde{G}_d^n} \partial_x \omega_0(\phi(x), d_x\phi \cdot T) J_T\phi(x) d\nu_x(T).\end{aligned}$$

We can also rewrite the third term as:

$$\begin{aligned}(3) &= \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \tilde{\gamma}(\phi, x, T) \; \mathrm{div}_T u(x) \; d\mu_0(x,T), \\ &\text{with} \;\; \tilde{\gamma}(\phi, x, T) = \omega_0(\phi(x), d_x\phi \cdot T) \; J_T\phi(x).\end{aligned}$$

As for the second term in (.1), for each (x,T) the integrand involves a linear combination (depending on ϕ) of the partial derivatives of u along the subspace T i.e. of the elements of the matrix $d_x u|_T \in \mathbb{R}^{n \times d}$. Thus, without attempting to specify this term

explicitly, we can in general write it as $\tilde{\beta}(\phi, x, T) d_x u|_T$ where $\tilde{B}$ is a continuous map from $\mathcal{B} \times \mathbb{R}^n \times \tilde{G}_d^n$ into $\mathcal{L}(\mathbb{R}^{n\times d}, \mathbb{R})$ giving us

$$(2) = \int_{\mathbb{R}^n \times \widetilde{G}_d^n} \tilde{B}(\phi, x, T) d_x u|_T \; d\mu_0(x, T).$$

The result of the theorem then follows by setting $\alpha(x) \doteq \tilde{\alpha}(\varphi_1^v, x)$, $\beta(x, T) \doteq \tilde{\beta}(\varphi_1^v, x, T)$ and $\gamma(x, T) = \tilde{\gamma}(\varphi_1^v, x, T)$. □

Reduced Hamiltonian equations:

Let V be the RKHS for a deformation group G_V which kernel is assumed to be radial scalar i.e. of the form $k_V(x, y) = \rho(|x - y|^2)\mathrm{id}_{\mathbb{R}^n}$. For convenience, let us denote $\rho(|x_j - x_i|^2)$, $\dot{\rho}(|x_j - x_i|^2)$, $\ddot{\rho}(|x_j - x_i|^2)$ and $\dddot{\rho}(|x_j - x_i|^2)$ by ρ_{ji}, $\dot{\rho}_{ji}$, $\ddot{\rho}_{ji}$, and $\dddot{\rho}_{ji}$. Then the reduced Hamiltonian equations (2.10) for the varifold diffeomorphic registration problem can be written more explicitly as follows:

$$\begin{cases}
\dot{x}_i & = \sum_{j=1}^N \left(\rho_{ji} p_k^x + 2 \sum_{k=1}^d \dot{\rho}_{ji} \langle x_j - x_i, u_j^{(k)} \rangle p_j^{u_k} \right) \\
\dot{u}_i^{(k)} & = \sum_{j=1}^N -2\dot{\rho}_{ji} \langle x_j - x_i, u_i^{(k)} \rangle p_j^x \\
& \quad - \sum_{j=1}^N \sum_{\ell=1}^d \left[4\ddot{\rho}_{ji} \langle x_j - x_i, u_i^{(k)} \rangle \langle x_j - x_i, u_j^{(\ell)} \rangle + 2\dot{\rho}_{ji} \langle u_i^{(k)}, u_j^{(\ell)} \rangle \right] p_j^{u_\ell} \\
\dot{p}_i^x & = \sum_{j=1}^N \Big\{ 2\dot{\rho}_{ji} \langle p_j^x, p_i^x \rangle + \sum_{\ell=1}^d 4\ddot{\rho}_{ji} \langle x_j - x_i, u_j^{(\ell)} \rangle \langle p_j^{u_\ell}, p_i^x \rangle \\
& \quad - \sum_{k=1}^d 4\ddot{\rho}_{ji} \langle x_j - x_i, u_i^{(k)} \rangle \langle p_j^x, p_i^{u_k} \rangle \\
& \quad - \sum_{k,\ell=1}^d 4\ddot{\rho}_{ji} \langle u_i^{(k)}, u_j^{(\ell)} \rangle \langle p_i^{u_k}, p_j^{u_\ell} \rangle \\
& \quad - \sum_{k,\ell=1}^d 8\dddot{\rho}_{ji} \langle x_j - x_i, u_i^{(k)} \rangle \langle x_j - x_i, u_j^{(\ell)} \rangle \langle p_i^{u_k}, p_j^{u_\ell} \rangle \Big\} (x_j - x_i) \\
& \quad + \sum_{j=1}^N \Big\{ \sum_{\ell=1}^d 2\dot{\rho}_{ji} \langle p_i^x, p_j^{u_\ell} \rangle - \sum_{k\ell=1}^d 4\ddot{\rho}_{ji} \langle x_j - x_i, u_i^{(k)} \rangle \langle p_i^{u_k}, p_j^{u_\ell} \rangle \Big\} u_j^{(\ell)} \\
& \quad - \sum_{j=1}^N \Big\{ \sum_{k=1}^d 2\dot{\rho}_{ji} \langle p_i^{u_k}, p_j^x \rangle + \sum_{k\ell=1}^d 4\ddot{\rho}_{ji} \langle x_j - x_i, u_j^{(\ell)} \rangle \langle p_i^{u_k}, p_j^{u_\ell} \rangle \Big\} u_j^{(k)} \\
\dot{p}_i^{u_k} & = \sum_{j=1}^N \Big\{ 2\dot{\rho}_{ji} \langle p_j^x, p_i^{u_k} \rangle - \sum_{\ell=1}^d 4\ddot{\rho}_{ji} \langle p_i^{u_k}, p_j^{u_\ell} \rangle \langle x_j - x_i, u_j^{\ell} \rangle \Big\} (x_j - x_i) \\
& \quad + \sum_{j=1}^N \sum_{\ell=1}^d 2\dot{\rho}_{ji} \langle p_i^{u_k}, p_j^{u_\ell} \rangle u_j^{(\ell)}
\end{cases}$$

Biographical sketch

Hsi-Wei Hsieh was born in Taipei, Taiwan in 1987. He received the B.A. degree in Economics from National Taiwan University in 2011, M.S. degree in Statistics from National Chiao Tung University in 2013, and enrolled in the Applied Mathematics and Statistics Ph.D. program at Johns Hopkins University in 2015. His research focuses on shape analysis via varifold representation, under the supervision of Dr. Nicolas Charon.